

FIGURE 1

CGGACGCGTGGGTGCGAGGCCGAAGGTGACCGGGGACCGAGCATTTTCAGATCTGCTCGGTAGA
 CCTGGTGCACCACCACC**ATG**TTGGCTGCAAGGCTGGTGTGTCTCCGGACACTACCTTCTAGG
 GTTTTCCACCCAGCTTTTCAACAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA
 ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAAACAAGAATTGGGATCCGGCGTGGGA
 GAACTGGCCAAGAAGTCAAAGAGGCAGCATTTGGAACCATCGATGGAAAAAATATTTAAATTT
 GATCAGATGGGAAGATGGTTTGTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA
 CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT
 ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAAACAGCT
 TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACTTCATGATGAGAGGCTCTTG
 GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC
 CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT
 GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCTCTTCTCATCAGAGCTGCATGGTACAC
 AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCAGTGAAAAGTTTCTGA
 ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCTCATTGGGATCTATG
 TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT
 AGTTCCTTTTTCAGCATGTTTCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT
 CACCAATGTATGGAGTTCAAAAATATGATCCCATTAAGTTCGATGCTGAGTATCTACATGGAT
 ACATTAAATATATTTATGCGAGTTGCAACTATGCTGGCAACTGGAGGCAACAGAAAGAAAT**AG**
AAGTGACTCAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAATGGGGCAGATATGC
 ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA
 TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT
 AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAATGTT
 TTGGTGAATGTGAAAACATAAGTTTGTGTGATGAGAATGTAAGTCTTTTTTCTACTTTAAAA
 TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTGTCATATTTTTTTGGAGT
 GCAGAATATTGTAATTAATGTCATAAGTGATTTGGAGCTTTGGTAAAGGGACCAGAGAGAAG
 GAGTCACCTGCAGTCTTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA
 GTGAGGAGCCAGTAAGAAACATCTGGGTATTTGGAAACAAGTGGTCATTGTTACATTCATTT
 GCTGAACTTAACAAAACGTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG
 CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGAAGTGTACAGAATGTTAATC
 ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTGAATGTTACAAAAGGAA
 ATAAGTTTAAACTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG
 AATACAAACAGTATACTCATG

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWFVAGGAAVGLGALCYGLGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLNFMMRGSWVTIGVTFAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLTIILGGPLLIIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGVLVFSMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCCGGTGGCGCCACGTCCGCCCCGTCTCCGCCTTCTGCAT
 CGCGGCTTCGGCGGCTTCCACCTAGACACCTAACAGTCGCGGAGCCGGCCGCGTCTGAGGG
 GGTCCGCACGGGAGTCGGGCGGTCTTGTGCATCTTGGCTACCTGTGGGTCTGAAG**ATG**TCGG
 ACATCGGAGACTGGTTCAGGAGCATCCCCGGCGATCACGCGCTATTGGTTCGCCGCCACCGTC
 GCCGTGCCCTTGGTCGGCAAACCTCGGCCTCATCAGCCCGGCCTACCTCTTCCTCTGGCCCCGA
 AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG
 GTCCAGGAACCTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA
 CTTGAAACAGGAGCTTTTGGATGGGAGGCCAGCAGACTATTATTTATGCTCCTCTTTAACTG
 GATTTGCATCGTGATTACTGGCTTAGCAATGGATATGCAGTTGCTGATGATTCCTCTGATCA
 TGTCAGTACTTTATGTCTGGGCCCAGCTGAACAGAGACATGATTGTATCATTTTTGGTTTGGGA
 ACACGATTTAAGGCCTGCTATTTACCCTGGGTATCCTTGGATTCAACTATATCATCGGAGG
 CTCGGTAATCAATGAGCTTATTGGAATCTGGTTGGACATCTTTATTTTTTCCTAATGTTCA
 GATACCCAATGGACTTGGGAGGAAGAAATTTCTATCCACACCTCAGTTTTTGTACCGCTGG
 CTGCCCAGTAGGAGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC
 TGCTGATCAGAATGGCGGAGGCGGGAGACACAACCTGGGGCCAGGGCTTTCGACTTGGAGACC
AGTGAAGGGGCGGCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCTCCAGTGCTGGGTG
 CACTTAACAACATGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC
 AGTACGAGACAAAGTTTCTTAAATCCCAGAAAAATATAAGTGTCCACAAGTTTCACGAT
 TCTCATTCAAGTCCTTACTGCTGTGAAGAACAAATACCAACTGTGCAAAATGCAAACTGAC
 TACATTTTTTGGTGTCTTCTCTTCTCCCCTTCCGTCTGAATAATGGGTTTTAGCGGGTCCCT
 AATCTGCTGGCATTGAGCTGGGGCTGGGTACCAAACCCCTTCCCAAAGGACCTTATCTCTT
 TCTTGCACACATGCCTCTCTCCCACCTTTCCCAAACCCCCACATTTGCAACTAGAAAAAGTTG
 CCCATAAAATTGCTCTGCCCTTGACAGGTTCTGTTATTTATTGACTTTTGCCAAGGCTGGTC
 ACAACAATCATATTCACGTTATTTTCCCCTTTTGGTGGCAGAACTGTTACCAATAGGGGGAG
 AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAAATTGCTTCGACTGACATCCGTTGTT
 AACCGTTTGGCACTCTTCAGATATTTTTTATAAAAAAAGTACCACTGAGTTCATGAGGGCCA
 CAGATTGGTTATTAATGAGATACGAGGGTTGGTGTGGGTGTTTGTTCCTGAGCTAAGTGA
 TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTAAACCATGGGGGATGCACCCC
 TTTGCGTTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTGCTTTGTGT
 TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTG
 ATTCCCATTTCATCTCATTCTGGATATGTGTTCAATTGAGTAAAGGAGGAGAGACCCCTCATA
 CGCTATTTAAATGTCACCTTTTTTGCCTATCCCCCGTTTTTTGGTTCATGTTTCAATTAATTGT
 GAGGAAGGCGCAGCTCCTCTCTGCACGTAGATCATTTTTTAAAGCTAATGTAAGCACATCTA
 AGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGGGTGTGTTA
 TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT
 TCGTAGGTGGGCTTTTCCCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCA
 TGGCTTTTTCACACAGTTATTTTATTTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGT
 ATTGAGTGGCTGTACACACTTTGAGGCAACTAAAAAGGCTTCAAACGTTTTGATCAGTTTCTT
 TTCAGGAAACATTGTGCTCTAACAGTATGACTATTCTTTCCCCCACTCTTAAACAGTGTGAT
 GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTTGTGTG
 TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA
 TGTTTTGTTGTTTCATCTGTGGCCACAATAAAGTTTACTTGTAAAATTTTAGAGGCCATTACT
 CCAATTATGTTGCACGTACACTATTGTACAGGCGTGGAGACTCATTGTATGTATAAGAATA
 TTTCTGACAGTGAGTGACCCGGAGTCTCTGGGTGACCCCTCTTACCAGTCAGCTGCCTGCGAG
 CAGTCATTTTTTCTTAAAGGTTTACAAGTATTTAGAACTTTTCAGTTCAGGGCAAAATGTTC
 ATGAAGTTATTCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT
 ATGTTTCTGGAATAATTTTACCAAACAAGCTATTTGAGTTTTGACTTGACAAGGCAAAACA
 TGACAGTGGATTCTCTTTACAAATGGAAAAAAAATCCTTATTTTGTATAAAGGACTTCCC
 TTTTTGTAACTAATCCTTTTATTGGTAAAAATTGTAAATTAAATGTGCAACTTG

FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF
PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIP
LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFL
MERYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL
GDQ

Transmembrane domain:

amino acids 98-116, 152-172

N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

Glycosaminoglycan attachment site.

amino acids 218-222

Downloaded from www.jstor.org

[The page contains faint, illegible markings or bleed-through from the reverse side.]

[The page contains faint, illegible markings or bleed-through from the reverse side.]

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTGGAATAATTCAAGAGAGAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAATCAAAGAGGAATATGACCAGGAA

FIGURE 8

GCGTGGTTTTTGTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTG
 TAGCTTCTCCACGTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG
 TCTCAGCTCTAGGATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTA AAAAAC
 AGTGGAATGAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATAC
 ATTCCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT
 CTGCCAATGAAGAAAACAAGTATGATTATCTTCCAACACTGTGAATGTGTGCTCAGAACTG
 GTGAAGCTAGTTTTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAG
 AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCTGCT
 TTCTTTATTTTCTGATAACTTGATTGTCTTCTATGTCCTGTCTATCTTCAACCAGCCATG
 GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTCAAGGATAGTGCTGAA
 GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCT
 TGACTGCCGGGACTAAACTTTACAGCACAACCTTGGCAGGACGTGGATTTCATCACGATGCC
 TTTTTCAGCCCTTCCAATTCCTGCCTTCTTTTCAGAAAGTGAGTGTCCAGAAAAGACAATTG
 TACAGCAAAGGAATGGACTTTTCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTTCAGTC
 ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT
 AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA
 GAACAGCAAACCTCTATTTCTTTGGCATTCTGTTTAATGGGCTGACTCTGGGCCTTCAGAGGA
 GTAACCGTGATCAGATTAAGAACTGTGGATTTTTTTATGGCCACAGTGCATTTTTCAGTAGCC
 CTTATTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA
 CATGTTCCATGTCTTGATGGCCCAGGTTACCACTGTCATTATCACAAACAGTGTCTGTCTGCTG
 TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGAAGCCCCATCAGTCTTCTCTCTATA
 TTTATTTATAATGCCAGCAAGCCTCAAGTTCGGAATACGCACCTAGGCAAGAAAGGATCCG
 AGATCTAAGTGGCAATCTTTGGGAGCGTTCAGTGGGGATGGAGAAGAACTAGAAAGACTTA
 CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTTCTAACTGGTACCCACATAGTTTGCA
 GCTCTCTTGAACCTTATTTTTCACATTTTCAGTGTTTGTAAATATTTATCTTTTCACTTTGATA
 AACCAGAAATGTTTCTAAATCCTAATATCTTTGCATATATCTAGCTACTCCCTAAATGGTT
 CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTCTAAAGAACTGATACAGGAGTAACA
 ATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT
 TTTCTTGGCCTTCAAGCTTCCAAAAACTTGTAATAATCATGTTAGCTATAGCTTGTATAT
 ACACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGT
 CTTCCCTTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGCCCTAGAGATAGT
 CATTTTGCAAGTAAAGAGCAACGGGACCCTTTCTAAAAACGTTGGTTGAAGGACCTAAATAC
 CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTGCTGAAGAAGCAGT
 TTCTCAGACACAACATCTCAGAATTTTAATTTTGTAGAAATTCATGGGAAATTGGATTTTTGT
 AATAATCTTTTGATGTTTTAAACATTGGTTCCCTAGTCACCATAGTTACCACTTGTATTTTA
 AGTCATTTAAACAAGCCACGGTGGGGCTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT
 GTCATTACTCCTGAATTATTACATTTTGGAGAATAAGAGGGCATTTTTATTTTATTAGTTACT
 AATTCAAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC
 CAGATTGTCAGTGAAGCTGATGCCTAGGAACCTTTAAAGGGATCCTTTCAAAGGATCACTT
 AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC
 AGTAATATATAAGTCACTTTACAGTGCTACTTCACACTTAAAGTGCATGGTATTTTTTCATG
 GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA
 AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG
 CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT
 CATTCTCAGAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA
 AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG
 GTGGAAATTTGTAATTAAATAATTATTAAACCT

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV
IFSNFSIITALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI
FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI
YNASKPQVPEYAPRQERIRDLSGNLWERSSSGDGEELERLTKPKSDESEDTF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293,
298-318, 328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

FIGURE 10

CGTGCCTGCGCAATGGGTGTCGGGTCCGCTTTTTCCCAATCCGGACGTAATCGTGGTTTTTG
TTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTGTAGCTTCTCCAC
GTATGGACCCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAG
GATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAAAAACAGTGGAATGGAA
AAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCACAATGTATACATTCTGCTAGG
TGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATTCTGCCAATGAAG
AAAACAAGTATGATTATCTTCCAACACTGTGAATGTGTGCTCAGAACTGGTGAAGCTAGTT
TTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATA
TGCTTCCTGGAAGGAATTCTCTGATTCATGAAGTGGTCCATTCTGCTTTCTTTATTTCC
TGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATGGCTGTTATCTTC
TCAAATTTTAGCATTATAACAACAGCTCTTCTATTTCAGGATAGTGCTGAAGAGGCGTCTAAA
CTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCTTGACTGCCGGGA
CTAAAACCTTA

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGGCCGGCTTGCGTAGCGCGCGGCGGCC
GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGCAGAGGAGCAT
CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCCGCGGGTCATGGCCAAAGGAGAAGGCGCCGAG
AGCGGCTCCGCGGCGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCCGCCCA
GGTGAAGAAAGAACCMAAAAGAAGAAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG
CACTTGGGGGAGCCCCCTACCAGGTGACGGGCTGTGCCCTGGGTTTCTTCCCTCAGATCTAC
CTATTGG**ATG**TGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC
CTGGGATGCCATCACAGACCCCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC
TGGGTCGCCTTATGCCCTGGATCATCTTCTCCACGCCCCCTGGCCGTCATTGCCTACTTCCTC
ATCTGGTTTCGTGCCCGACTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT
CTTTGAAACAATGGTCACTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTTCATCAGCA
ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC
AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG
ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATAACATGGCACCACCTTCACAC
AGGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTG
TGCTGTCACTCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTG
AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATAACATCAAACCTT
ATTACTGGCTTCCTCTTCACCTCCTTGGCTTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTT
TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT
CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT
GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA
CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC
TACCTTGGTCCATGCTGCCTGATGTCAATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT
GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTTCACCAAGTTTGCTCTGGAGTGT
ACTGGGCATTTCTACCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC
CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG
CTGGGCCTGCTGCTCTTCAAAATGTACCCCATTGATGAGGAGAGGCGGCGGCAGAATAAGAA
GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG
AGCTGGCTAGCATCCTCT**AG**GGGCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG
AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA
CTGAAGACTCAAGGAGGTGGCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG
TGGCCTCCTGCCTCCCCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA
TGCCAAGGACTGATCGGGCCTAGCCCGGAACACTAATGTAGAAACCTTTTTTTTTTACAGAGCC
TAATTAATAACTTAATGACTGTGTACATAGCAATGTGTGTGTATGTATATGTCTGTGAGCTA
TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

FIGURE 12

MWLRWALSLPPSSCLWAEPMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG
SCPTSHTARPIGTCTFSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL
GTAIQGQIVGQADTPCFQDFNSSTVASQSANHTHGTTSHRETQKAYLLAAGVIVCIYIICAV
ILILGVREQREPYEAAQQSEPIAYFRGLRLVMSHGPIYIKLITGFLFTSLAFMLVEGNFVLFCT
YTLGFRNEFQNLLLAIMLSATLTIPWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI
ITYAVAVAAGISVAAFLLPWSMLPDVIDDFHLKQPHFHGTAPIFFSFYVFFTKFASGVSLG
ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKAL
QALRDEASSSGCSETDSTELASIL

12/330 12/330 12/330

FIGURE 13

GGGAAACGCAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTTA
T TACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACCTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGTTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCAT

FIGURE 14

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT
ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAA**ATGT**
GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT
TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT
CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG
CAGTTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA
GAGAACGTTATCATCAAATTAACAAGGCTGGCCTTGTAAGTGGAAATACTGAGTTGTTTAGG
ACTTTCTATTGTGGCAAACCTTCAGAAAACAACCCCTTTTGTCTGCACATGTAAGTGGAGCTG
TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATG
CAGCCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG
AGTAAGTGCACTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG
ATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT
ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGA
TTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCCTCTATGACACTG
CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCAGAGATATT**TGAT**GAAAGGAT
AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAGTTGCTTA
TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA
ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA
TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACTATG

FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFI¹FSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI
AAVLCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG
AVLTFGMGSLYMFVQ²TILSYQM³QPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLT⁴YIRDFQKISLRVEANLHGLTLYD
TAPCPINNERTRL⁵LSRDI

123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

FIGURE 16

CGGACGCTTGGGCNGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGTGCCTGATGCCGAGT
TCCGTCTCTCGGGTCTTTTCCTGGTCCCAGGCAAAGCGGAGCGGAGATCCTCAAACGGCCTA
GTGCTTCGCGCTTCCGGAGAAAATCAGCGGTCTAATTAATTCCTCTGGTTTGTTGAAGCAGT
TACCAAGAATCTTCAACCCTTTCCCACAAAAGCTAATTGAGTACACGTTCCCTGTTGAGTACA
CGTTCCTGTTGATTTACAAAAGGTGCAGGTATGAGCAGGTCTGAAGACTAACATTTTGTGAA
GTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCT
TCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTTTCATACATTACTGCAGTAACACT
CCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTANC

T05111 5332550

FIGURE 17

CCACGCGTCCGCCCCCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCGCGGG
CCGGGGTGCGGAGCCGACATGCGCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC
CTTCGCCTTGTACTTGCTGTGACGCGACTGCCCCGCGGGCGGAGACTGGGCTCCACCGAGG
AGGCTGGAGGCAGGTGCGTGTGGTTCCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG
GTCTTCGAGAGTACCGGAAGGAGCACCCAGGCCCTACGTGTTCTTGCTCTTCTGCGGCGCCTA
CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCTGAATGTTTTAGCTGGTGCCT
TGTTTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTTGACCTCGGTGGGTGCCACATGC
TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCTACTTTCTGATAAAGT
GGCCCTGCTGCAGAGAAAGGTGGAGGAGAACAGAAACAGCTTGTTTTTTTTCTTATTGTTTT
TGAGACTTTTCCCCATGACACCAAACCTGGTTCTTGAACCTCTCGGCCCAATTCTGAACATT
CCCATCGTGCAGTTCTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTCATCTGTGT
GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG
TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAAATTCCTGGAACCCCTATTAAAAAATTT
AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA
CACTGATCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA
TGTGGTCTCTAAAGCCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG
CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT
TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTTGT
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTCT
ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAAT
ACAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC
AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT
CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLLPRGRRLLGSTEELAGGRSLWFPSDLAELRELSEVLREYR
KEHQAYVFLFLFCGAYLYKQGFAIPGSSFLNVLAGALFGPWLGLLLCCVLTSGATCCYLLSS
IFGKQLVVSYPDPKVALLRKVEENRNSLFFFLFLRLFPMTPNWFLNLSAPILNIPIVQFF
FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLLIKKFSQKHLQ
LNETSTANHIHSRKDT

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31

Sugar transport proteins.

amino acids 162-172

FIGURE 19

CCGAGGCGGGAGGAGCCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATTTT
 CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAA
 TAGGAAAATAACTTGGGATTTTATATTGGAAGACATGGATCTTGCTGCCAACGAGATCAGCA
 TTTATGACAACTTTCAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATG
 TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC
 CCCCCGCGAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT
 TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT
 GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA
 GAAGTACATGTCAGAAAATAAGGGAGTTCCTCTGCATGGGGGTGATGAAGACAGACCCTTTC
 CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCTGCC
 AACTGCACTGGCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA
 ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC
 AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAG
 TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCTTATCCATGGAGGAGACCTCTGAA
 CAGATCACAAATGTTACGTGAGCTTTTTCTGTTTTCACTCACCTGCCATTTCCAAAAGATG
 CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG
 ATGCCTGACCTATTTATCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCA
 GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCG
 ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT
 GGAACCGCTTCTCAGAACTGTAGGAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA
 AAACCAGGTTGAAAGGGGAAAAATAAAAACAAAAACGATGAAACTGCAAAAA

FIGURE 20

MDLAANEISIIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY
KVLATLGLILLTAYFVIQPFSPPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL
HGGDEDRFPDFDPWWTNDCEQNESEPI PANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT
GKPLLEEEIQHFLCQYPEATEGFSEGGFAKWWRCFPERWFFPYPWRRPLNRSQMLRELFV
FTHLPFPKDasLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRHRCQSVAMP
IEPGDIGYVDTTHWKVYVIARGVQPLVICDGTAFSEL

FIGURE 21

CCACGGTGTCCGTTCTTCGCCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCGAGGGGCG
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTCAGTTGTGATC
AAGGGACACGTGGTTTCCGAAC TGCCAGCTCAGAATAGGAAAATAACTTGGGATTTTATATT
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAAACTTTCAGAGACTGTTG
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCAGAGAAGGCAATTGAAAAATTTATC
AGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCCGCAGTATCCTCTCCTTATAGT
TGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

FIGURE 22

CCCACGCGTCCGCCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG
 CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGTGCTCCTGGCTGGTG
 CCACTGGTGCACGCTGCTAGACCGTGCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCC
 CTCCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTTGTTGCCA
 CACCCGAATGGCGCCACTTTCATCGACAAACAGGTACAGCCAACC**ATG**TCCAGTTTCGAAATG
 GACACGTATGCTAAGAGCCACGACCTTATGTTCAGGTTTCTGGAATGCCTGCTATGACATGCT
 TATGAGCAGTGGGCAGCGGCGCCAGTGGGAGCGCGCCAGAGTCGTCGGGCCCTTCCAGGAGC
 TGGTGCTGGAACCTGCGCAGAGGCGGCGCCCTGGAGGGGCTACGCTACACGGCAGTGTG
 AAGCAGCAGGCAACGCAGCACTCCATGGCCCTGTGCTGACTGGGGGGCGCTGTGGCGCCAGCT
 CGCCAGCCCATGTGGGGCCTGGGCGCTGAGGGACACTCCCATCCCCCGCTGGAAACTGTCCA
 GCGCCGAGACATATTACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCCTCAC
 CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCCCTGACACCCACCGAGGAGGC
 CTCCTGCTTCTGGCAGTGACCAAAGAGGCCAAAGTGAGCAGCCCAACCCAGTTGCTGCAGG
 AGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGAACTG
 GATGAGCAGCGTGAGAAGCTGGTGCTGTGCGCCGAGTGCCAGCTGGTGACGGTAGTGGCCGT
 GGTCCCAGGGCTGCTGGAGGTACACACACAGAATGTATACTTCTACGATGGCAGCACTGAGC
 GCGTGGAAACCGAGGAGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCAGCTGCGTGAG
 GTCCACCTGCGGCGTTTCAACCTGCGCGCTTGACACTTGAGCTCTTCTTTATCGATCAGGC
 CAACTACTTCTCAACTTCCATGCAAGGTGGGACGACCCAGTCTCATCTCCTAGCAGAGA
 CTCCGAGACCCAGCCTGGCCCCATCCCACCCCATACCCAGGTACGGAACAGGTGTACTCG
 TGGCTCCTGCGCCTACGGCCCCCTCTCAAGGCTACCTAAGCAGCCGCTCCCCCAGGAGAT
 GCTGCGTGCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACTTCGAGTACT
 TGATGCAACTCAACACCATTTGCGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTT
 CCTGGGCTCTGCAGGACTACGTGTCCCCAACCCCTGGACCTCAGCAACCCAGCCGTCTTCCG
 GGACCTGTCTAAGCCCATCGGTGTGGTGAACCCCAAGCATGCCAGCTCGTGAGGGAGAAGT
 ATGAAAGCTTTGAGGACCCAGCAGGGACCATTGACAAGTTCACCTATGGCACCCACTACTCC
 AATGCAGCAGGCGTGATGCACTACCTCATCCGCGTGAGGCCCTTACCTCCCTGCACGTCCA
 GCTGCAAAGTGGCCGCTTTGACTGCTCCGACCGGCAGTTCCACTCGGTGGCGGCAGCTGGC
 AGGCAGCCTGGAGAGCCCTGCCGATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATT
 TCAGCAACTTTGGGCAGACTCCCTGTGAGTGTGTAAGGAGCCACATCCAACCTCGGCTCTCA
 GCTGAGGAAGCAGCCCATCGCCTTGACGCGCTGGACACTAAGTACCTAGCATCTTCCAGCA
 CCTGGACGAACTCAAGGCATTCTTCGAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA
 CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTTCAGCTTCAGCAAGAGACCC
 ACCATGGGCAGCCACAAGACGCAGCGACTGTGAGTGGCCCGTGAGTGGCAGGCAGTGGTGT
 GAGTGGACAAGCACTGGCAGTGGCCCCGGATGGAAAGCTGCTATTACGCGGTGGCCACTGGG
 ATGGCAGCCTGCGGGTGACTGCACTACCCCGTGGAAGCTGTTGAGCCAGCTCAGCTGCCAC
 CTGATGTAGTAACCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCG
 GGACACCAGCTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGAGTAGGCCTGGCAC
 CAAAGCCTGTGCAGGTCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACT
 GAACTTGACATGGCTGTGTCTGGATCTGAGGATGGAAGTGTGATCATAACACTGTACGCCG
 CGGACAGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACC
 TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCTGGGGCC
 CAGGTACCTACTCCTTGACCTGTATTAGTCAATGGGAAGTTGCGGGCTTCACTGCCCT
 GGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCAGT
 GCGCCCTGCACATCCTCCAACATAACACACTGCTCCCGGCCGCGCCTCCCTTGCCCATGAAG
 GTGGCCATCCGCAGCGTGGCCGTGACCAAGGAGCGCAGCCACGTGCTGGTGGGCCCTGGAGGA
 TGGCAAGCTCATCGTGGTGGTTCGCGGGGCGAGCCCTCTGAGGTGCGCAGCAGCCAGTTTCGCGC
 GGAAGCTGTGGCGTCTCGCGGCGCATCTCCAGGTGTCTCGGAGAGAGCAGGAATACAAC
 CTAAGTGGGCGCGCT**TGA**ACTTGGCCAGTCCGGCTGCTCGGGCCCCCGCCCGGCAGGCGCTG
 GCGGGGAGGCCCCGCCCAGAAGTCGGCGGGAACACCCCGGGGTGGGCAGCCAGGGGGTGA
 GCGGGGGCCACCCTGCCAGCTCAGGGATTGGCGGGCGATGTTACCCCTCAGGGATTGGCG
 GGCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGGCCGCCCTGAGGGCCAGCACTGGCGTCT

FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL
 RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN
 HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPELLQEDQLGEDELAEELETP
 MEAAELDEQREKLVLSAECQLVTVVAVVPGLLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP
 LAQLREVHLRRFNLRSALELFFIDQANYFLNFPCKVGTTTPVSSPSQTPRPQPGPIPPHTQV
 RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYNDL
 SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH
 YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADVKELEIP
 EFFYFPDFLENQNGFDLGCLQLTNEKVGDVVLPWASSPEDFIQQHRQALESEYVSAHLHEW
 IDLIFGYKQRGPAEEALNVFYCTYEGAVDLDHVTDERERKALEGIISNFGQTPCQLLKEP
 HPTRLSAEEAAHRLARLDTNPSPIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF
 SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWDGSLRVTALPRGKLL
 SQLSCHLDVVTCLALDTCGIYILISGRDTCMVWRLLHQGGLSVGLAPKPVQVLYGHGAAVS
 CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSA
 WERPGAQVTYSLHLYSVNGKLRLASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA
 PPLPMKVAIRSVAVTKERSHVLVVGLEDGKLIVVVAGQPSEVRSSQFARKLWRSSRRISQVSS
 GETEYNPTAR

N-glycosylation site.

amino acids 677-681

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 985-989

Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

N-myristoylation site.

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697,
 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840,
 873-879, 912-918, 954-960

FIGURE 24

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCC
 CACGGCCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT
 CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC
 TGGACCCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT
 CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC
 GCACACTCCGTTACCACACTGGGTCAATTGGCATTGAGGCCCTCATCCTGACCCTTGTGCAG
 ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC
 CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC
 TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA
 AATGCGTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGA
 CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCTGTCTTCTTTT
 TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCACCTCAACTATTAC
 TGGCTGCCCATCATGACCTCCATCCTGGGGGCTATGTCATCGCCAGCGGCTTCTTCAGCGT
 TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCTGGAAGACCTGGAGCGGAACAACG
 GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC
 GAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGGCCCTGATCCAGGACTGC
 ACCCCACCCACCCTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT
 AAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG
 AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC
 GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC
 TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCGGGAGGCAGAGGTTGCAGTGAGCCGAGA
 TCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAACAA
 AAAGATTTTATTAAAGATATTTTGTAACTC

FIGURE 25

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLGF
WTLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGNFCVSAK
NAFMLLMRNIVRVVVLDKVTDLLFFGKLLVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKN
EAPPDNKKRKK

25/330

FIGURE 26

GAGTCTTGACCGCCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT
 GGCT**ATG**TTCGTGTCCGATTTCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC
 TTCTCTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTTC
 CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAAACTGC
 ATTTCTTGAGCATAAAGAACAGTTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG
 ACCTATTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGG
 CCAGTCAATGTCTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGA
 TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT
 CAGGAAATGACAGTGATGGGTGAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA
 GTGGAGCAAACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT
 CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG
 CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC
 CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCTGCAGCG
 CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA
 CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC
 AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA
 GCGGCTCCAGGAGTTCCTTGACAGATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC
 AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA
 TTTGGGATGAAGGACATGCGCGTGCAGACTTTCAGCATTTCATTTTGGGTTCAGCACAAAGTT
 TCTGGCCAGCGACGTGGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT
 CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG
 TACCATGGCCTGGAACCTGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC
 CTTTGCACCAACCTCGTCATCTCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC
 TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA
 AGTCCTTTGTGTGTTGACAAAGAACCGGCGCTGCAAACCTGCTGCCCCTGGTGATGGCTGCC
 CCCCTGAGCATGGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTC
 GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGAAAGCACCAGCTCCCGGA
 TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT
 CTGGACGCACTTATTTCCCTCCTGTCC**TAG**GAATTTGATTCTTCCAGAATGACCTTCTTATT
 TATGTAACCTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT
 TTTTATTAAATAAAAATGCTTATTTTAGGAAA

FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETA
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAIEDIFRDEEEDEEHSGNDS DGSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLRH
VSRHNHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKR
LQEF LADMGLPLKQVKQKFQAMD ISLKENLREMIEESANKFGMKDMRVQTF SIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLRSNLDKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP
LSMEHGTVTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNHFDLSVIELKAEDRSKFL
DALISLLS

404121525360

FIGURE 28

GTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGTGGCTATGNTCGTGTCCGATTTCCGCA
AAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTTCTCTTCGTGGCCTCGGANGTGGAT
GCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGACCANGTGCAATATANGCT
GGTTCCAGTTTCTGGGTGGCAAGAACTTGAAACTGCATTTCTTGAGCATAAAGAACAGTTTC
ATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTATTGGATATTCTTCAACCTGAT
GAAGACACTATATTCTTTGTGTGTGACACCCATAGGCCAGTCAATGTTGTCAATGTATACAA
CGATACCC

FIGURE 29

CAGGAACCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAAC
 CACGAAGGGACGATACCAGAAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTG
 GCTGACTTTTGGCTATAGAAAAAAGAAAGGAACGAAAAGAGACAGTTTTTTTTTGGAAAGCTAA
 GTCTTCCCTTTATCGAGTCAAGAAACCCCCCTTCTTGAGCTATTTACAGCTTTTAAACAATT
 GAGTAAAGTACGCTCCGGTCACC**ATG**GTGACAGCCGCCCTGGGTCCCCTCTGGGCAGCGCTC
 CTGCTCTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGC
 CAGCGGCTGCCAACGGTGCTGTGACTCTGAGGACCCCTGGATCCTGCCCATGTATCCTCAG
 CCTCTTCTCCGGCCGCCACGCCCTGCCTGAGATCAGACCCTACATTAATATCACCATC
 CTGAAGGGTGACAAAGGGGACCCAGGCCCAATGGGCCTGCCAGGGTACATGGGCAGGGAGGG
 TCCCCAAGGGGAGCCTGGCCCTCAGGGCAGCAAGGGTGACAAGGGGGAGATGGGCAGCCCCG
 GCGCCCCGTGCCAGAAGCGCTTCTTCGCCTTCTCAGTGGGCGCAAGACGGCCCTGCACAGC
 GGCGAGGACTTCCAGACGCTGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGA
 CATGGCGACCGGCCAGTTTGCTGCTCCCTGCGTGGCATCTACTTCTTACGCTCAATGTGC
 ACAGCTGGAATTACAAGGAGACGTACGTGCACATTATGCATAACCGAAAGAGGCTGTCTATC
 CTGTACGCGCAGCCCAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGC
 CTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAACGCCATCTACAGCA
 ACGACTTCGACACCTACATCACCTTCAGCGGCCACCTCATCAAGGCCGAGGACGACT**GA**GGG
 CCTCTGGGCCACCCCTCCCGGCTGGAGAGCTCAGGTGCTGGTCCCCTCCCCTGCAGGGCTCAG
 TTTGCACTGCTGTGAAGCAGGAAGGCCAGGGAGGTCCCCGGGGACCTGGCATTTCTGGGGAGA
 CCTTGCTTCTATCTTGGCTGCCATCATCCCTCCAGCCTATTTCTGCTCCTCTCTTCTCTCT
 TGGACCTATTTTAAAGAGCTTGCTAACCTAAATATTCTAGAACTTTCCCAGCCTCGTAGCCC
 AGCACTTCTCAAACCTTGGAATGCATGCGAATCACCCGGGGTTCGTGTTAAATGCAGATTCT
 GACTCAGCAGGTCTGAGTGGGTCCAGGATTCTGTGTTTCTCATATGTTCCCTGGGTGATGCTG
 ATGGGGTCAGTCTATGAACCACACTGGAGCAACCAGGTTCTAGGACTTTCTCAATATTCTAG
 TACTTTCTGAACATTCTGGAATCCTCCCCACATTCTAGAATTCTCCCAACATTTTTTTTTCT
 TGAGACAGAGTCTTGCTCTGTTGCCAGGCTAGAGTGCAGTGGTGCAATCTCAGTTCAGTGC
 AACCTCTGCCTCCCGGGTTCAGCGATTCTTCTGCCTCAGCCTCCCTAGTGGCTGGGATTAC
 AGGCGCCTGCTACCATGCCTGGCTAATTTTTGTATTTTTTAGTAGAGATGGGGTTTACCATA
 TTGGCCAGGCTGGTCTTGAACCTCCTGACTTCAGGTGACCCACCCGCTCGGCCCTCTCAAAT
 GCTGGGATTACAGGTGTGAGCCACCGTGCCTGGCCAATTCCAACATTCTTAAATTCTCTCAT
 CCCTCCAGGGCTCCCCGTGCTATGTTCTCTTTACCCCTTCCCCCTCTTCTCTTGCTCAGGCC
 TGCACCACTGCAGCCACCGTTCAATTTATTCATTCAATAAACTGAGCACTCACTCTGTGCT
 GGGTCCCCGGGAAGGGTGAGGGGGTACAGACAGGCCCTGCCCCTGCCCTCAGTGAAGTGGCCA
 GTCCAGCCCAGGCGGGGAGAGATGTGTACATAGGTTTTAAAGCAGACCCAGAGCTCATGGGG
 GCCTGTGTTCTGGGTGTTCAAGGTGCTGCTGGTCCATTACCCACTGCTCCCCAAGGCTGG
 TGGGACGGGGTCCCGGTGGCAGGGGCAGGTATCTCCTTCCCGTTCCTCATCCACCTGCCAG
 TGCTCATCGTTACAGCAAACCCAGGGGGCCTTGGCCAGGTCAAGGGTTCTGTGAGGAGAGG
 ACCCAGGAGTGTGGGGGCATTTGGGGGGTGAAGTGGCCCCGAAGAATGGAACCCACACCCA
 TAGCTCTCCCCACAGCTGATACGGCATCCTGCGAGAAGACCTGCCCTCCTCACTGGGATCCC
 CTTCTGCTCCTCCCAGGGCTCTGCCAGGGCCTTGCTCAGTCCCTTCCACCAAAGTCATCT
 GAACTTCCGTTTCCCCAGGGCCTCCAGCTGCCCTCAGACACTGATGTCTGTCCCAGGTGCT
 CTCTGCCCTCATGCCCTCTCACCGGCCAGTGCCCGGACTCTCCAGGCTTTATCAAGGTG
 CTAAGGCCCGGGTGGGCAGCTCCTCGTCTCAGAGCCCTCCTCCGGCCTGGTGCTGCCTTTAC
 AAACACCTGCAGGAGAAGGGCCACGGAAGCCCCAGGCTTTAGAGCCCTCAGCAGGTCTGGGG
 AGCTAGAGCAAAGGAGGGACCTCAGGCCTTCCGTTTCTTCTTCCAGGGTGGGGTGGCCTGGT
 GTTCCCCTAGCCTTCCAAACCCAGGTGGCCTGCCCTTCTCCCAGAGGGAGGCGGCCTCCGC
 CCATTGGTGCTCATGCAGACTCTGGGGCTGAGGTGCCCCGGGGGGTGATCTCTGGTGCTCAC
 AGCCGAGGGAGCCGTGGCTCCATGGCCAGATGACGGAACAGGGTCTGACCAAGTGCCAGGA
 AGACCTGTGCTATAAACCCCTGCCTGATCCTGCCCTGCCTGACCCCGCCACGCCCTGCC
 GTCCAGCATGATTAAAGAATGCTGTCTCCTCTTGAAAAA

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSSGRPH
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRF
FAFSVGRKTALHSGEDFQTLLEFVFNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET
YVHIMHNQKEAVILYAQPSEERSIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDITYIT
FSGHLIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

Clq domain proteins.

amino acids 144-178, 78-111 and 84-117

MCSRVP L L L L L L L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N
G I T M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R
R L E R L Y L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P
G I L D T A N V E A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G
N T R I A Q L R P E D L A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G
P W V R E S H V T L A S P E E T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S
S S L A P T W L S P T A P A T E A P S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E
G F T G L Y C E S Q M G Q G T R P S P T P V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L
T Y R N L S G P D K R L V T L R L P A S L A E Y T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A
V H S N H A P V T Q A R E G N L P L L I A P A L A A V L L A A L A A V G A A Y C V R R R G R A M A A A A Q D K G Q V G P G A G
P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E C E V P L M G F P G P G L Q S P L H A K P Y I

FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT
CTTCATCATTTCATATGAGGAAATAAGTGGTAAATCCTTGGAAATACA**ATG**AGACTCATCAG
AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG
AAGAAAGGGAAGTATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG
ACCCAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTTCAAG
TTTTCAATTCTGTCTCCAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCACAGCTGG
ATCTCAAAACCTTTGAATTCAACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG
AAGAGTGTAACCTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTAATGACTT
TGACACCATGCTATCTGTGAGGAAGCTGGCAACATGTCACACCTGGAAATCCTAGGTTTGA
GTGGGGCAAAAATACAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTC
TTCTTAGGATTCAGAACTCTTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC
AAAACCTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTCTTTTGGCTGATGGAATCA
AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA
ATGCAACGAAATCTTAGTTTAGAAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA
TTTACTCTGGGACGACCTTTTCTTATCTTACAATTTGTTTGGCATAACATCAGTGGAACT
TTCAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTAC
TCAATACTGTAATGAGAACTATAAAATTTGGAGCATGTACATTTTCAAGAGTGTTTTACATTCA
ACAGGATAAAATCTATTTGCTTTTGGACCAAAATGGACATAGAAAACCTGACAATATCAAATG
CACAAATGCCACACATGCTTTTCCGAAATATCTACGAAATTCGAATATTTAAATTTTGCC
AATAATATCTTAACAGACGAGTTGTTTAAAAGAACTATCCAACCTGCCTCACTTGAAAACCT
CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC
CCTTGGAACTTGGATCTGAGTCAAAATCTATTACAACATAAAAATGATGAAAATTTGCTCA
TGGCCAGAAACTGTGGTCAATATGAATCTGTCTACATAAATTTGTCTGATTCTGTCTCAG
GTGCTTGCCCAAAAGTATTCAAATACTTGACCTAAATAATAACCAATCCAACTGTACCTA
AAGAGACTATTCATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTTTCTAACTGAT
CTCCCTGGATGCAGTCATTTTCACTAGACTTTTCACTTCTGAACATTGAAATGAACCTTCATTCT
CAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAACTCTAAATGCGGGAAGAA
ATCCATTCGGTGTACCTGTGAATTAATAAAATTTTCACTCAGCTTGAAACATATTCAGAGGTC
ATGATGGTTGGATGGTCAGATTCATACACCTGTGAATACCCTTTAAACCTAAGGGGAAGTAG
GTTAAAAGACGTTTCTCTCCACGAATTATCTTGCAACACAGCTCTGTTGATTGTCACCATTG
TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCCC
TGGTATCTCAGGATGCTAGGTCAATGCACACAACATGGCACAGGGTTAGGAAAACAACCCA
AGAACAACCTCAAGAGAAATGTCCGATTCCACGCATTTATTTTATACAGTGAACATGATTCTC
TGTGGGTGAAGAATGAATGATCCCCAATCTAGAGAAGGAAGATGGTTCTATCTTGATTTGC
CTTTATGAAAGCTACTTTGACCTGGCAAAAGCATTAGTGAAAATATTGTAAGCTTCATTGA
GAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCAACTTTGTCCAGAATGAGTGGTGCCATT
ATGAATTCTACTTTGCCCACCACAATCTCTCCATGAAAATTTCTGATCATATAATTCTTATC
TTACTGGAACCCATTCCATTCTATTGCATTTCCACCAGGTATCATAAACTGAAAGCTCTCCT
GGAAAAAAGCATACTTTGGAATGGCCCAAGGATAGGCGTAAATGTGGGCTTTTCTGGGCAA
ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAACTGCAGACA
TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT
ATAAATCCCACAGTCTTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA
CAACCTTTATGATGGCAATTTGACAATATTTATTAATAAAAAATGGTTATTTCCCTTCATA
TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG
GCTTATGAAAAAGGTGTTTATCCCAGGATTGTTTATAATCATGAAAAATGTGGCCAGGTGC
AGTGGCTCACTCTTGTAATCCCAGCACTATGGGAGGCCAAGGTGGGTGACCCACGAGGTCAA
GAGATGGAGACCATCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTA
GCTGGGCGTGATGGTGACGCCTGTAGTCCAGCTACTTTGGGAGGCTGAGGCAGGAGAATCG
CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACTCCAGCCTGGT
GACAGAGCGAGACTCCATCTCAAAAAAAGAAAAAAGAAAAAATGGAAAACATCC
TCATGGCCACAAAATAAGGTCTAATTCAATAAATTTATAGTACATTAATGTAATATAATATTA
CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCTGGTATGGAAAAACATATTAATAT
GTTATAAACTATTAGGTTGGTGCAAACTAATTGTGGTTTTTGCCATTGAAATGGCATTGAA
ATAAAGTGTAAGAAATCTATACCAGATGTAGTAACAGTGGTTGGGTCTGGGAGGTTGGA
TTACAGGGAGCATTTGATTTCTATGTTGTGTATTTCTATAATGTTTGAATTGTTTAGAATGA
ATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQDLKTFFFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTkMDIENLTISNAQMPHMLFPNYPTKfQYLNfANNILtDELfKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLsqNLLQHkNDENCsWPETVvNMNLSYNKLS
DSVFRCLPKSIQILDlnnnQIQTVPKETIHLMALRELNIafNfLTDLPGCSHFsrLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVGLAVAFcCLHFDLPWYLRMLGQCTQTWHRV
RKTTQEQlKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHhNLFHENSdHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKCGLFWANLRAAINVNVLATREMYELQTfTELNEESRGSTISLM
RTDCL

FIGURE 35

GGGGGCTTTCTTGGGCTTGGCTGCTTGGAAACCTGCCTCCAAGGACCGGCCTCGGAGGGGTGCGCGGGAAACG
 GAGGGAAGAAGGAAGGGCGGGCGGGCCCCCTGCGCCCGCCCCGCGCCTCTGCGCGCCCTGTCCGCCCCGGC
 CCAGCCCAGCCAGCCCGGGGCGGTACACGCGCAGCCAGCCGGCGCCCTCCCGCGCCCAAGCGCGCCGCT
 CTGCTGTGCCCTGCGCCCTTGGCCCGCGCCAGCTTCTGCGCCCGCAGCCCGCCGGCGCCCCGGTGACCGTGA
 CCCTGCCCTGGGCGCGGGGCGGAGCAGGCATGCTCCCGCCCCGGGACCGCTACCCAGCGCTGGCCCTGGTGCTC
 CTGGCAGTGACCCTGGCCGGGGTCCGAGCCAGGGCGCAGCCCTCGAGGACCCTGATTATTACGGGCAGGAGAT
 CTGGAGCCGGGAGCCCTACTACGCGCGCCCGAGCCCGAGCTCGAGACCTTCTCTCCGCCGCTGCCTGCGGGGC
 CCGGGGAGGAGTGGGAGCGGCGCCCGCAGGAGCCAGGCCGCCAAGAGGGCCACCAAGCCCAAGAAAGCTCCC
 AAGAGGGAGAAGTCGGCTCCGGAGCCGCTCCACCAGGTAACACAGCAACAAAAAGTTATGAGAACCAAGAG
 CTCTGAGAAGGC TGCCAACGATGATCACAGTGTCCGTGTGGCCCGTGAAGATGTCAGAGAGAGTTGCCACCTC
 TTGGTCTGGAACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGGCA
 CATCGAGGGAGACTCAACATCCAGGCGGGCATTAAATGAAATGATTTTATGACGGAGCGTGGTGCGCGGGAAG
 AAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATTCACTGGTGTCTCACTCAAGGGA
 GGAATCCCTCTGGCTGAGTGAAGTGGTGAATCCTATAAGGTGATGGTGAGCAATGACAGCCACACGTGGGTC
 ACTGTTAAGAATGGATCTGGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCT
 ACCCGTCCCATGGTGGCCCGCTACATCCGCATAAACCTCAGTCCCTGGTTTGATAATGGGAGCATCTGCATGA
 GAATGGAGATCCTGGGCTGCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCACT
 GATGACCTGGATTTTAAAGCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTGAATGAAATGTGTCC
 CAATATCACCAGAATTTACAACATTGGAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATC
 ACCCTGGGGAGCATGAAGTCGGTGAGCCCGAGTTCCACTACATCGCGGGGGCCACGGCAATGAGGTGCTGGGC
 CGGGAGCTGCTGCTGCTGCTGGTGCAGTTTCGTGTGTGAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCT
 GGTGGAGGAGACGCGGATTACGTCTCCCTCCCTCAACCCGATGGCTACGAGAAGCCCTACGAGAAGGGGCT
 CGGAGCTGGGAGGCTGGTCCCTGGGACGCTGGACCACGATGGAATTGACATCAACAACAACCTTCCCTGATTTA
 AACACGCTGCTCTGGGAGGCGAGGATCGACAGAATGTCCCGAGGAAAGTTCCCAATCACTATATTGCAATCCC
 TGAGTGGTTTCTGTGCGGAAATGCCACCGTGGCTGCCGAGACCAGAGCAGTCATAGCCTGGATGGAAAAATCC
 CTTTTGTGCTGGGCGGCAACCTGCAGGGCGGCGAGCTGGTGGTGGCGTATCCCTACGACCTGGTGCGGTCCCC
 TGGAAGACGCAGGAACACACCCCCACCCCCGATGACCAGTGTTCGCTGGCTGGCCTACTCCTATGCCTCCAC
 ACACCGCCTCATGACAGACGCCCCGAGGAGGTGTGCCACACGGAGGACTTCCAGAAGGAGGAGGCACTGTCA
 ATGGGGCCTCCTGGCACACCGTCTGCTGGAAGTCTGAACGATTTTCACTACCTTCATACAACTGCTTCGAACTG
 TCCATCTACGTGGGCTGTGATAAATACCCACATGAGAGCCAGCTGCCCCGAGGAGTGGGAGAATAACCGGGAATC
 TCTGATCGTGTTTCATGGAGCAGGTTTCATCGTGGCATTAAGGCTTGGTGAGAGATTACATGGAAAAGGAATCC
 CAAACGCCATTATCTCCGTAGAAGGCATTAAACATGACATCCGAACAGCCAACGATGGGGATTACTGGCGCCTC
 CTGAACCTTGAGAGTATGTGGTACAGCAAAGGCCGAAGGTTTCACTGCATCCACCAAGAACTGTATGGTTGG
 CTATGACATGGGGGCCACAAGGTGTGACTTCACTTACAAACCAACATGGCCAGGATCCGAGAGATCATGG
 AGAAGTTTGGGAAGCAGCCGTCAGCCTGCCAGCCAGGCGGCTGAAGCTGCGGGGGCGGAAGAGACGACAGCGT
 GGGTGAACCTCCTGGGCCCTTGAGACTCGTCTGGGACCCATGCAAAATTAACCAACCTGGTAGTAGCTCCATAG
 TGGACTCACTCACTGTTGTTTCTCTGTAATTCAAGAAGTGCCTGGAAGAGAGGGTGCATTGTGAGGCAGGTCC
 CAAAAGGGAAGGCTGGAGGCTGAGGCTGTTTTCTTTTCTTTGTTTCCATTTATCCAAATAACTTGGACAGAGCA
 GCAGAGAAAAGCTGATGGGAGTGAGAGAACTCAGCAAGCCAACCTGGGAATCAGAGAGAGAAGGAGAAGGAGG
 GAGCCTGTCCGTTAGAGCCTCTGGCTGCATAGAAAAGGATTCTGGTGCTTCCCTGTTTGGCGTGGCAGCAAGG
 GTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAGCATTTCCCGAGCTGGGCTGTCCCAATGTTACCA
 TTTGAGATGCTCCAGGCGTCTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTACAAATAA
 ATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACATCAGTGAGCCTCTTGAATCTGTTTAGTCTCCT
 TTTTCAACAAAGGAGTGTGTTGAGAAAAGGAGAGAGAGGCTGAGATCATTAGGAGTTTGTGGGCGAGCAAGCA
 TGGAGCTTCTTGACAAATCTGGGTCCATAAACACCCCCAAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTT
 CCCCAGGTAGGGAGAGCCAGAGGTGCCAGCCTTCTGAAGGGCCAGAAAATTTAGCCTGGATCTCCTCTTTTAC
 CTGCTAGGACTGGAAGAGCCAGAAGTGGGGTGGCCTGAAGCCCTCTCTCTGCTTGAGGTATTGCCCTGTGTG
 GAATTGAGTGCTCATGGGTGGCCTCATATCAGCCTGGGAGTTATTTTGTATATGTAGAATGCCAGATCTTCCA
 GATTAGGCTAAATGTAATGAAACCTCTTAGGATTATCTGTGGAGCATCAGTTTGGGAAGAATTATTGAATTAT
 CTTGCAAGAAAAAGTATGTCTCACTTTTTGTTAATGTTGCTGCCTCATTGACCTGGGAAAAATGAAAAAAA
 AATAAAGCAAATGGTAAGACCTTAAAAA

FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYYARPEPELETFSPPLP
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPVMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPEGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG
SELGGWSLGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRL
NPGEYVVTAKEGFTASTKNCMVGYDMGATRCDFTLSTNMARIREIMEKFGKQPVSLPARR
LKLGRKRQRG

FIGURE 37

CTAAGAGGACAAGATGAGGCCCGGCTCTCATTTCTCCTAGCCCTTCTGTTCTTCTTGGCCAAGCTGCAGGGG
 ATTTGGGGGATGTGGGACCTCCAATTCAGCCCCGGCTTCAGCTCTTTCCAGGTGTTGACTCCAGCTCCAGC
 TTCAGCTCCAGCTCCAGGTCCGGCTCCAGCTCCAGCCGAGCTTAGGCAGCGGAGGTTCTGTGTCCCAGTTGTT
 TTCCAATTTACCCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC
 CCGTGGACAGAGTGAACGCTTGGAAATTCACAGCTCATGTTCTTTCTCAGAAGTTTGAGAAAGAACTTCTAAA
 GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAACTGTTAAACCTAACTGTCCGAATGACATCAT
 GGAGAAGGATACCATTTCTTACACTGAACCTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAC
 TGGTCATACAGCTGAAGGAGAGTTTTGGTGAAGCTCAGAAATGTTGACCAGCTGGAGGTGGAGATAAGAAAT
 ATGACTCTCTTGGTAGAGAAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTCGCCGAGAAATCGTGGC
 TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCCTCTAAAGATCAAAACACCCCTGTGTCACCCCTCTCCCACTC
 CAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGTTT
 TCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTCCCAGCATCCAAACAAAGGACTGTATTGGGTGGCGCC
 ATTTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTTGCTATGTATATAA
 ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC
 ATGTACAACACCGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACCTCTCCCTAA
 TGCTGCCCTATAATAACCGCTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAAATG
 GATTGTGGGTTATTTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAACTCAATGACACCACACTT
 CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCTTCATGGTATGTGGGGTTCT
 GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGGAAGAGG
 GCAAACCTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGCAGAGCATTAACTATAACCCCTTTTGACCAGAAA
 CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTCTGTCTTGCAGAAGCCCCAGTAAAGCTGTTA
 GGAGTTAGGGTGAAAGAGAAAAATGTTTGTGAAAAAATAGTCTTCTCCACTTACTTAGATATCTGCAGGGGTGT
 CTAAAAGTGTGTTTCATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT
 TGATTTGGTGAGTTCTCTTGGGAATCATCTGCCTCTTCAGGCGCATTTTGCAATAAAGTCTGTCTAGGGTGGGA
 TTGTGAGAGGTCTAGGGGCACTGTGGGCCCTAGTGAAGCCTACTGTGAGGAGGCTTCACCTAGAAGCCTTAAATTA
 GGAATTAAGGAACCTTAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC
 CTCATCCATGTAGCACCCTAATTCTTCCATGCCTGGAAGAAACCTGGGGACTTAGTTAGGTAGATTAATATCT
 GGAGCTCCTCGAGGGACCAATCTCCAACCTTTTTTCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG
 CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCCTTTT
 ATGCATTAAATTGTACATGGCAAATAAATCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTTCTCTC
 ATTGTCCACCTTACTAAAAGTCAGTAGAATCTTCTACCTCATAACTTCTTCCAAAGGCAGCTCAGAAGATTAG
 AACCAGACTTACTAACCAATTCCACCCCCACCAACCCCTTCTACTGCCTACTTTAAAAAAATTAATAGTTTT
 CTATGGAACCTGATCTAAGATTAGAAAAATTAATTTCTTTAATTTTATTATGGACTTTTATTTACATGACTCTA
 AGACTATAAGAAAAATCTGATGGCAGTGACAAAGTGCTAGCATTTATTGTTATCTAATAAAGACCTTGGAGCATA
 TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTGCCTTTGTTTAAGCCTGGAACCTGTAAGAAAT
 GAAAATTTAATTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT
 TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT
 TCCTTTGATGTTCAAGTCCTAGTCTATAGGATTGGCAGTTTAAATGCTTTACTCCCCCTTTTAAATAAATGAT
 TAAATGTGCTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 38

MRPGLSFLLLALLFFLLGQAAGDLGDVGPPIPSPGFSSFPGVDSSSSSFSSSSSRSGSSSSSRSLGS
GGSVSQLFNFTGSVDDRGTQCQSVSLPDTTFPVDRVERLEFTAHVLSQKFEEKELSKVREYV
QLISVYEKKLLNLTVRIDIMEKDTISYTELDFELIKVEVKEMEKLVIQLKESFGGSSEIVDQ
LEVEIRNMTLLVEKLETLDKNNVLAIRREIVALKTKLKECEASKDQNTPVVHPPPTPGSCGH
GGVVNISKPSVQNLNWRGFSYLYGAWGRDYSPOHPNKGlyWVAPLNTDGRLLLEYRlyNTLD
DLLLYINARELRITYGQSGTAVYNNNMVYNMYNTGNIARVNLTTNTIAVTQTLPNAAYNnr
FSYANVAWQDIDFAVDENGLWVIYSTEASTGNMVISKLNDTTLQVLNTWYTKQYKPSASNAF
MVCGLYATRtMnTRTEEIfYYYDTNTGKEGKLdIVMHKMqEKVQSINYNPFdQKLYVYNDG
YLLNYDLsVLQKPQ

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCCTCCAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA
TGTACAACACCGGGNATATTGCCAGAGTTAACCTGACC

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQVLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMYLTSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVTYSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVVWKISSSWIGI
VLYVWTLVAPLVLTNRDFD

FIGURE 42

GCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATAACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGTCTATGCCGATGCTGTCCTAGTGGAACAANTCCACTGTAAGTAGATTGATCTA
TGCACTTTTCTTGCTTGTGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAG
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

42/330

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTGTTCATGGCGAGCTGGATAC
 CANGTTTGTGTGGAAGTGCCCCGTGTTTGNATGCCGATGCTGTCCTAGTGGAACAANTCC
 ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGTAAT
 GTTGATACCAGGAATGGAAGAACAAGTGAATAAGATTCCCTGGATTTTGTGAGAATGAGAAAG
 GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
 GCTANGTTCTATNTTCTTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
 AGCTGCAGTGCACAATGGATTTTGGTTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG
 GGGC

FIGURE 45

GCTGTCCTTAGTGGAAACAANTCCAACCTTGTAACCTGGATTGATCTATGCACTTTTTTCCTTG
CTTGTTGGAGTATGTGTAGCTTTGTGTAATGTTGTTCCCAGGATTGGANGAACAACTGAATA
AGATTCCTGGATTTTTGTGAGAATGAGAAAGGTGTTGTCCCCTTGTAACATTTTTGGTTGGC
TATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTACT
AATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCT
TTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGCATTCTTCATTCCAGAAGGAACTTTT
ACAACTGTGTGGTTTTATGTAGGCATGGCAGGTGCCTTTTGTTCATCCTCATACAACCTAGT
CTTACTTATTGATTTTGCACATTCATGGAATGAATCGTGGGTGAAAAAATGGAAGAAGGGA
ACTCGAGATGTTGGTATGCAGCCTTGTTATCAGCTACAGCTCTGAATTATCTGCTGTCTTTA
GTTGCTATCGTCCTGTTCTTTGTCTACTACACTCATCCAGCCAGTTGTTTCAGAAAACAAGGC
GTTCATCAGTGTCAACATGCTCCTCTGCGTTGGTGCTTCTGTAATG

FIGURE 46

CTCGGGCGCGCACAGGCAGCTCGGTTTGCCTGCGATTGAGCTGCGGGTCGCGGCCGGCGCCGGCCTCTCCAAT
 GGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTCGGCAAAGGCAGTCGAGTGTTCGAGACCGGGGCGAG
 TCCTGTGAAAGCAGATAAAAGAAAACATTTATTAACGTGTCATTACGAGGGGAGCGCCCCGGCCGGGGCTGTGCG
 ACTCCCCGCGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTAC
 GTCGTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCAGATATTTGATTTATTAGCGATGCC
 CCCTGGTTTGTGTGTTACGCACACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCC
 TGGGCGAATCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTCGAATCTGCGAGTG
 AAGAGGGACGAGGGAAAAGAAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAAGACACCAGAT
 CAGCAAAAAAAGAAGATGGGGCCCCCGAGCCTCGTGCTGTGCTTGTGTCGCGCAACTGTGTTCTCCCTGCTGGG
 TGGAAAGCTCGGCCTTCCCTGTCGCACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCA
 ACATCATCCTGCTGCTGACGGACGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGC
 ATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCCCTCAGCTCCTC
 CATCCTCACTGGCAAGTACGTCCACAACCACAACACTACACCAACAATGAGAAGTGTCTCCGCCCCCTGGC
 AGGCACAGCAGCAGAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAG
 TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTGCGACTCCTTAAAAACTCCCG
 CTTTTATAACTACACGCTGTGTCGGAACGGGTGAAAGAGAAGCAGGCTCCGACTACTCCAAGGATTACCTCA
 CAGACCTCATACCAATGACAGCGTGAGCTTCTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCTCTC
 ATGGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTACAGCCCCACAATATTCAGCCTCTTCCCAAACGC
 ATCTCAGCACATCAGCCGAGCTACAACCTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC
 CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGCCCTCATGTGCGTGGAC
 GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGC
 CGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGG
 TCCCGTTCTACGTGAGGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG
 GCCCCACCATCCTGGACATTGCAGGCCTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCT
 GGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG
 AGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCCCAGTAC
 CAGCGTGTGAAGGACCTGTGTGTCAGCGTGCTGAGTACCAGACGGCGTGAGCAGCTGGGACAGAAGTGGCAGTG
 TGTGGAGGACGCCACGGGGAAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC
 TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTC
 AGCCTGGCCGGACGCCGAAAAAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCGAGTGTCTCCATCCG
 CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA
 CCAAGCGGCACTGGCCAGGGGGCCCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGC
 CTTTCCGACTACTCAGCCGCCAACCCCATTAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCA
 GTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGA
 CCCTGCAGAACAAAATTAAGAACCTGAGGGAAAGTCCGAGGTACCTGAAGAAAAGCGGCCAGAAGAATGTGAC
 TGTACAAAATCAGCTACCACACCCAGCACAAAGGCCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTAG
 GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGC
 TCAAGCGCCTGCAGAACACGACACGTGCAGCATGCCAGGCCTCACGTGCTTACCCACGACAACCAGCACTGG
 CAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGCTGACACAGCGCCAACAATAACAGTACTGGTGCAT
 GAGGACCATCAATGAGACTCACAATTTCTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCA
 ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCTCAACCAGCTACACGTACAG
 CTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGGACTCGAAACATGGACCTGGATGGAGG
 AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG
 GACAACCTGTGGGAAGGCTGGGAAGGTAAAGAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA
 CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGC
 AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC
 ATTTTTGCCCTGCTTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTCTGATCAAAAAGTCA
 ACCACTAACCTCCCCAGAAGCTCACAAAGGAAAACGGAGAGAGCGAGCGAGAGAGATTTCTTGGAAATTTT
 TCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAGCAGTCCTGTTCTAAATCCTCTTATTCTT
 TTGGTTTGTGACAAAGAAGGAAGTAAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG
 TTTGACAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCTGGTTGCCCTGTGAAGAAA
 CTGCTTCATTGTATATGTGACTATTTACATGTAATCAACATGGGAACCTTTAGGGGAACCTAATAAGAAAT
 CCCAATTTTCAGGAGTGGTGGTGTCAATAACGCTCTGTGGCCAGTGTAAAAGAAAAA

FIGURE 47

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQ
VMNKTRRIMEQGGAHFINAFVTTMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHE
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFNPASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN
FLPKYQRVKDLQCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSVVAIEVDGRVYHVGLGDAAQ
PRNLTKRHWP GAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRCLKHRGSSL
HPFRKGLQEKDKVLLREQKRKKLRLKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTYWC MRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

FIGURE 48

AACAAAGTTCAGTGAAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA
GCTGCTGGGCAGAGAGGGGACTGTCCGGCTCCCAGATGCTGGGCCTCCTGGGGAGCACAGCCC
TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCCTGCTGCTGCTGCTGCTGCTGGCCACC
TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGGAAA
CCGAGTCCGCCGGGCCCAGCCTTGGCCCTTCCGGCGGGGGGCCACCTGGGAATCTTTCACC
ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC
CCCCGCCACACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCCGCCA
CGCTCGCTGAGGCTGCTGTGCGCCGGTGCCTGTGGACAGCAGCTGCCCCTGCCCTCCCATCTG
TTCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG
GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC
AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGTGAAGGGTTTGGGGAGTGGAGAGCAAGG
GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGAAGTCCCAGTGAAGTGGAGAGCAAGG
ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGGCCGTCTTGACCTCCTC
ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT
GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT
AAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 51

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGlyGETGR
LFTESCSISPklRSIAVYYDNPHMVPPDKCrcAVGSILSEGEESpSPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATRrvHPALDTYIKERKLCAYPRLEIYQEDQIHfMCPLAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSaATLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEgPLGESRLDPGTEPLGTTKWLWEPTAPEK
GKE

Protein

FIGURE 52

CCGCGGGAACGCTGTCCTGGCTGCCGCCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCT
GCCCCGCGCCAGTCA**ATG**ACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCT
GCTGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGA
CCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGA
GACACGCTTCACATACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCT
GACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGA
GTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCCTTCTCACTTGGCCTAT
GGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCT
GATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAG
GGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAAT
AGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAAT**TA**
ATAAATAATAAATTTTAAAAAACTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 53

MTLRPSLLPLHLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGLTLHI
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS
KKKLKEEKRNKSXXX

0997-1501
532660

FIGURE 54

CCCGGGAACGTGTTCCCTGGCTGCCGCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCTGC
CCCGCGCCCAGTCATGACCCTGCGCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCTGC
TGCTCAGTGCGGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGACC
CTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGAGA
CACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCTGA
CCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGAGT
CTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTATGG
AAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCTGA
TTTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAGGG
ATGGCCATGGTGCCACCCTCCTGGGCCTCATTTGGGTATCACCTATACAGAAAGGCCAATAGA
CCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAATAATA
AATAATAAATTTTAAAAAACTTA

FIGURE 55

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAAGTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAA
CAAGAGCAAAAAGAAATAATAATAATAAATTTAAAAAACTTAAAA

FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAG
TGAGGCGGGCCGGCGCGCGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG
ACCTGAAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG
GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT
TATCATAGATGCAGCTGTTATTTATCCACCATGAAAGATTTCAACCACTCATACCATGCCT
GTGGTGTTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA
GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTTGGCTTTTCGTTGG
TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG
CTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT
TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAGTTGAACACATCTGAT
TTCCACAGCACAACAGCCCTGCATGGGTTTGTTTGTTTTTTACTGCTCACTCCCAACCTT
TTGTAATGCCATTTTCTAAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT
AAAATCACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACTT
ATAAAATGTTAGAGGAACTTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATT
TGTA AAAAATAAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA
TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT
CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTTAGTTTTTTAAAATATTCCGTGG
TCAA AATTCTTCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT
GTA ACTGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCAT
GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCTCTCAT
CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC
CACATCCACCACTG

FIGURE 58

TTCTTGGCTAAAATCGGGGGAGTGAGGCGGGCCGGCGGGCGGACACCGGGCTCCGGAACC
ACTGCACGACGGGGCTGGACTGACCTGAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATG
CTCAGAATGCATTTGACTGGGGGGAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTAC
TATTTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGAT
TTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCATAGCCTTCCTAATGATTAATGC
AGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTG
CTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGG
ATTCTTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATT
TTTCCAGAATGCCTTCATCTTTTTTTGGAGGGCTGGTTTTTTAAGTTTGGC

FIGURE 59

TGGACGGACCTGAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATGTTTCAGAATGCATGAC
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTACAGGCTG
GTGGATTATCATAGATGCAGNTGTTATTTATCCCACCATGAAAGATTTCAACCANTCATACC
ATGCCTGTGGTGTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT
CGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTTGGAGGTT
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC
ATCTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

FIGURE 60

GGACACCGGGTTCCGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAATGTTTGGATT
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTGGTGGATTATCATAGATGCAGCTGTTATT
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTATAGCAACCATAGC
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT
GTTTGGGTCAAACAGGTGNTCGCATTGCGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATTT
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAGACATAGTAT
ACCCTGGAATTNCTNTATTTTTTCCAGAATGCC

60/330

FIGURE 62

GGGAGGCTGTGNCCGTTTTGTTTTNTTGGCTAAAATCGGGGGAGTGAGGCGGCCCGGC GCGG
CGNGACACCGGGTTCCGGGAACCATTGCACGACGGGGTGGACTGACCTGAAAAAAATGTTTG
GATTTNTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATT
GCTTCCATTGCTGCTGGTGTACTATTTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGT
TATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCA
TAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAA
GGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGG
ATNTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAG
TATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATNTTTTTTTGGAGGGCTG

FIGURE 63

CGACGCCGGCGT**GATG**TGGCTTCCGCTGGTGCTGCTCCTGGCTGTGCTGCTGCTGGCCGTCC
TCTGCAAAGTTTACTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC
AAACGGCCCCCAGCGCCCCTGGTAACTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC
TTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG
GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCCTGGTGCTGGAACAACAT
ACCAAGGCAGGGGGCTGCTGTCTATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAAT
CCATTACATTGGGCGTATGGAAGAGGGCAGCATTGGCCGTTTTATCTTGGACCAGATCACTG
AAGGGCAGCTGGACTGGGCTCCCCTGTCTCTCTTTTGACATCATGGTACTGGAAGGGCCC
AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAGAAAGCCTACATTCAAGGCCTCAAGGA
GAAGTTTCCACAGGAGGAAGCTATCATTTGACAAGTATATAAAGCTGGTTAAGGTGGTATCCA
GTGGAGCCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTGCCCCTGGTTTCAGCTCCTCGAC
AGGTGTGGGCTGCTGACTCGTTTTCTCTCCATTCTTCAAGCATCCACCCAGAGCCTGGCTGA
GGTCTGTCAGCAGCTGGGGGCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCA
CTTACGGTGTCAACCCCAACCCACAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTAC
ATGAAAGGAGGCTTTTATCCCCGAGGGGGTTCCAGTGAAATTGCCTTCCACACCATCCCTGT
GATTTCAGCGGGCTGGGGGCGCTGTCTCACAAAGGCCACTGTGCAGAGTGTGTTGCTGGACT
CAGCTGGGAAAGCCTGTGGTGTGAGTGTGAAGAAGGGGCATGAGCTGGTGAACATCTATTGC
CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGC
CCGCTGCCTGCCAGGTGTGAAGCAGCAACTGGGGACGGTGCAGGCCGGCTTAGGCATGACCT
CTGTTTTCTATCTGCCTGCGAGGCACCAAGGAAGACCTGCATCTGCCGTCCACCAACTACTAT
GTTTACTATGACACGGACATGGACCAGGCGATGGAGCGCTACGTCTCCATGCCAGGGGAAGA
GGCTGCGGAACACATCCCTCTTCTCTTCTCGCTTTCCCATCAGCCAAAGATCCGACCTGGG
AGGACCGATTCCCAGGCCGGTCCACCATGATCATGCTCATAACCACTGCCTACGAGTGGTTT
GAGGAGTGGCAGGCGGAGCTGAAGGGAAGCGGGGCAGTGACTATGAGACCTTCAAAAACCTC
CTTTGTGGAAGCCTCTATGTGAGTGGTCTGAACTGTTCCCAAGCTGGAGGGGAAGGTGG
AGAGTGTGACTGCAGGATCCCCACTACCAACCAAGTTCTATCTGGCTGCTCCCCGAGGTGCC
TGCTACGGGGCTGACCATGACCTGGGCGCGCTGCACCCCTTGTGTGATGGCCTCCTTGAGGGC
CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTCACTGTGGACTGGTCCG
GGGCCCTGCAAGGTGCCCTGCTGTGCAGCAGCGCCATCCTGAAGCGGAACCTGTACTCAGAC
CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAAT**TAG**TTCCATCAGGAGG
AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG
CATTAGTTCTTGCACGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG
TTTAAATCACAAATCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA
TCTTTACGCCTTTTATAACATGCCATCCCTACTAATAGGATATTGACTTGGATAGCTTGATG
TCTCATGACGAGCGGCGCTCTGCATCCCTCACCCATGCCTCCTAACTCAGTGATCAAAGCGA
ATATTCCATCTGTGGATAGAACCCCTGGCAGTGTGTGCTCAGCTCAACCTGGTGGGTTCACTTC
TGTCTTGAGGCTTCTGCTCTCATTCAATTTAGTGCTACGCTGCACAGTTCTACACTGTCAAGG
GAAAAGGGAGACTAATGAGGCTTAACCTCAAAACCTGGGCGTGGTTTTGGTTGCCATTCCATA
GGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTCAAGTGGCTCTTCAGGGGACAGGAAAT
GCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGTTAGTA
GGGTGCATGTGAGATGATCATATCCAATTCATATGGAAGTCCCGGGTCTGTCTTCTTATCA
TCGGGGTGGCAGCTGGTTCTCAATGTGCCAGCAGGGACTCAGTACCTGAGCCTCAATCAAGC
CTTATCCACCAAATACACAGGGAAGGGTGATGCAGGGAAGGGTGACATCAGGAGTCAGGGCA
TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCCAGCCAAGGG
CACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAGGGGAAGTCACATCAGAAAAGGGA
AAGCCACGGAATGTGTGTGAAGCCCAGAAATGGCATTGTGAGTTAATTAGCACATGTGAGGG
TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAATGACTTTTCAGTTATGTCTTTG
GTATCAGACATACGAAAGGTCTCTTTGTAGTTCGTGTTAATGTAACATTAATAAATTTATTG
ATTCCATTGCTTTAAAAA

FIGURE 64

MWLPLVLLLLAVLLLLAVLCKVYLGLFSGSSPNPFSEDVKRPPAPLVTDKEARKKVLKQAFSAN
QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLFDTGIHYIG
RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ
EEAIIIDKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ
LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA
GGAVLTKATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP
GVKQQLGTVRPGGMTSVFICLRGTEKEDLHLPSTNYYVYYDTMDQAMERYVSMFREEAAEH
IPLLEFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEWQAELKGKRGSDYETFKN SFVEA
SMSVVLKLFPPQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI
PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLDSRIRAQKKKN

FIGURE 65

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCTA
 GGGGTTGGCACCGGCCCGAGAGGAGG**ATG**CGGGTCCGGATAGGGCTGACGCTGCTGCTGTG
 TGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGGATGAAGAAGGCAGCCAGGATGAATCCT
 TAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTAAAGGACCATACTACTGCAGGCAGA
 GTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGTCTGAATTAGAATCCTCTATTCAAGA
 AGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTACAGAAGATATCAGCTTTCTAG
 AGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTACGGAAACCAGCTTTG
 ACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCTTTTCTAGATAA
 GGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTGCTACAACCT
 ATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCTAAGAGA
 CGGCAGATGCAGGAAGCAGAAATGATGTATCAAAGTGAATGAAAATCCTTAATGGAAGCAA
 TAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
 CCAAAGCCCTGGAGAGAGTGTGATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATC
 CAGGCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGC
 TCTTGGCTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTAT
 ATTATACATTTGGAGCTCTTGGGGGCAATCTAATAGCCACATGGTTTTGGTAAGTAGACTT
TAGTGGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTT
 TCAGCTTTTCATGATCCAGATTTGCTTGTATTAAGACCAAATATTCAGTTGAACTTCCTTCAA
 ATTCTTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACA
 ATTTTTCTTTAAAATGATTAGTTTGGCTGATTGCCCCATAAAAAGAGAGATCTGATAAATGGC
 TCTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTACATTAGATTATCATAA
 TTTTAAAAATTTTTCTTTAGTTTTTCAAATTTTGTAAATGGTGGCTATAGAAAAACAACAT
 GAAATATTATACAATATTTTGCAACAATGCCCTAAGAATTGTTAAATTCATGGAGTTATTT
 GTGCAGAATGACTCCAGAGAGCTCTACTTTCTGTTTTTTACTTTTCATGATTGGCTGTCTTC
 CCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTCTCATTTTCC
 CTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACTCATTTTTTAATAAA
 ATTATGTCTAAGATTAAAAA
 AA
 AA

FIGURE 66

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLDSKTTLTSDSVKDHTTAGRVVAGQIFLD
SEESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHG
EPCHFPPFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARRQMGEAEMM
YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMF EK
LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLI AHMVLVSRL

FIGURE 67

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCT
GCCTCCCTGCCTCTGGCC**ATG**GCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCT
GTCAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGG
CTCAACTCTCCTGCACGCTCAGCCCCCAGCACGTCACCATCAGGGACTACGGTGTGTCCTGG
TACCAGCAGCGGGCAGGCAGTGCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCA
CCACCGGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCACAATGCCT
GTGTCCTCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGC
TACGGCTTTAGTCCC**TAG**GGGTGGGGTGTGAGATGGGTGCCTCCCCTCTGCCTCCCATTTCT
GCCCCTGACCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGG
TTAATAATATTCAACATGTCAACAAC

68/330

FIGURE 68

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPQVAQLSCTLSPQHVTIRDYGVSWYQQRAG
SAPRYLLYYRSEEDHHRPADIPDRFSAKDEAHNACVLTISPVPEDDADYYCSVGYGFS

Protein

FIGURE 69

GCCGCCCCGCCCCGAGACCGGGCCCCGGGGGCGCGGGGCGGGCGGGATGCGGGCGCCCCGGGGCGG
 CGATGACCGCGGAGCGCACGCCGCGGGCCCCGGCCCTGACCCCCGCGCCCCGCCCCGCTGAGCCC
 CCGGCCGAGGTCCGGACAGGCCGAG**ATG**ACGCCGAGCCCCCTGTTGCTGCTCCTGCTGCCGC
 CGCTGCTGCTGGGGGCCTTCCCACCGGCCGCCGCCCGAGGCCCCCCAAAGATGGCGGAC
 AAGGTGGTCCCACGGCAGGTGGCCCCGCTGGGCGCACTGTGCGGCTGCAGTGGCCAGTGGGA
 GGGGGACCCGCCGCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA
 GCCGCTTCCGCGTGCTGCCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGC
 GTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCCTCGTCGT
 GCTGGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCGACAGCTCCTCTGGGGGTCAAG
 AGGACCCCGCCAGCCAGCAGTGGGCACGACCGCGCTTCACACAGCCCTCCAAGATGAGGCGC
 CGGGTGATCGCACGGCCCCGTGGGTAGCTCCGTGCGGCTCAAGTGCGTGGCCAGCGGGCACCC
 TCGGCCCCGACATCACGTGGATGAAGGACGACCAGGCCCTTGACGCGCCAGAGGCCGCTGAGC
 CCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAACCTGCGGCCGGAGGACAGCGGCAATAC
 ACCTGCCGCGTGTGCAACCGCGCGGGCGCCATCAAGCCACCTACAAGGTGGATGTGATCCA
 CGGGACCCGTTCCAAGCCCGTGCTCACAGGCACGCACCCCGTGAACACGACGGTGGACTTCG
 GGGGGACCACGTCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG
 AAGCGCGTGGAGTACGGCGCCGAGGGCCGCCACAACCTCCACCATCGATGTGGGCGGCCAGAA
 GTTTGTGGTGCTGCCACGGGTGACGTGTGGTGCGGGCCGACGGCTCCTACCTCAATAAGC
 TGCTCATCACCCGTGCCCGCCAGGACGATGCGGGCATGTACATCTGCCTTGGCGCCAACACC
 ATGGGCTACAGCTTCCGCAGCGCCTTCTCACCGTGCTGCCAGACCCAAAACCGCCAGGGCC
 ACCTGTGGCCTCCTCGTCTCGGCCACTAGCCTGCCGTGGCCCGTGGTCATCGGCATCCCAG
 CCGGCGCTGTCTTCATCCTGGGCACCCTGCTCCTGTGGCTTTGCCAGGCCCAGAAGAAGCCG
 TGCACCCCGCGCCTGCCCTTCCCCTGCCTGGGCACCGCCCGCGGGGACGGCCCCGCGACCC
 GAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCCCTCAGCGCTGGCCCTGGTGTGGGCTGT
 GTGAGGAGCATGGGTCTCCGGCAGCCCCCAGCACTTACTGGGCCAGGCCAGTTGCTGGC
 CCTAAGTTGTACCCCAAACCTCTACACAGACATCCACACACACACACACACACTCTCACAC
 ACACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTATCAGTGCT**TAG**ACGGCACCCGT
 ATCTGCAGTGGGCACGGGGGGGCCGGCCAGACAGGCAGACTGGGAGGATGGAGGACGGAGCT
 GCAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCAGGCAGTCTGTGTG
 TGAGGCATAGCCCCCTGGACACACACACACAGACACACACACTACCTGGATGCATGTATGCAC
 ACACATGCGCGCACACGTGCTCCCTGAAGGCACACGTACGCACACGCACATGCACAGATATG
 CCGCCTGGGCACACAGATAAGCTGCCCAAATGCACGCACACGCACAGAGACATGCCAGAACA
 TACAAGGACATGCTGCCTGAACATAACACACGCACACCCATGCGCAGATGTGCTGCCTGGACA
 CACACACACACACAGATATGCTGTCTGGACACACAGATAATGCTGCCTTGACACACACATGCACGG
 ATATTGCCTGGACACACACACACACACACGCGTGCACAGATATGCTGTCTGGACACGCACAC
 ACATGCAGATATGCTGCCTGGACACACACTTCCAGACACACGTGCACAGGCGCAGATATGCT
 GCCTGGACACACGCAGATATGCTGTCTAGTCACACACACACGCAGACATGCTGTCCGGACAC
 ACACACGCATGCACAGATATGCTGTCCGGACACACACACGCACGCAGATATGCTGCCTGGAC
 ACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGATATTGCCTGGACACACACA
 TGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACACG
 CACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGACACACATGCACACACAGGT
 GCAGATATGCTGCCTGGACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGA
 TATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA
 TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA
 CACACATGCACACACAGGTGCAGATATGCTGCCTGGACACACGCAGACTGACGTGCTTTTGG
 GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTT
 CCTGCTCCACCGTCACTCCCCCAACTCTGCCCGCCTCTGTCCCCGCCTCAGTCCCCGCCTC
 CATCCCCGCCTCTGTCCCCGGCCTTGGCGGCTATTTTTGCCACCTGCCTTGGGTGCCCAGG
 AGTCCCCCTACTGCTGTGGGCTGGGGTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGCTGGAG
 CCCATGGCTAGTGGCTCATCCCCAGTGCATTCTCCCCCTGACACAGAGAAGGGGCCTTGGTA
 TTTATATTTAAGAAATGAAGATAATATTAATAATGATGGAAGGAAGACTGGGTTGCAGGGAC
 TGTGGTCTCTCCTGGGGCCCCGGGACCCGCTGGTCTTTTCAGCCATGCTGATGCCACACCCC
 TGGAGGCCACACACCCCCCACCCTGCTGCTGGTGGTGGCCCCAGATCTCTGTAATTTTA
 TGTAGAGTTTGGAGCTGAAGCCCCGTATATTTAATTTATTTTGTAAACACAAAA

FIGURE 71

CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAACTTCCCAGGGGACCGCATTCCAGAGTC
 AGTGACTCTGTGAAGCACCCACATCTACCTCTTGCCACGTTCCCACGGGCTTGGGGGAAAG**ATG**GTGGGGACCA
 AGGCCCTGGGTGTTCTCCTTCCCTGGTCTGGAAGTCACATCTGTGTTGGGGAGACAGACGATGCTCACCCAGTCA
 GTAAGAAGAGTCCAGCCTGGGAAGAAGAACCCAGCATCTTTGCCAAGCCTGCCGACACCCCTGGAGAGCCCTGG
 TGAGTGGACAACATGGTTCAACATCGACTACCCAGGCGGGAAGGGCGACTATGAGCGGCTGGACGCCATTGCT
 TCTACTATGGGGACCGTGTATGTGCCCGTCCCTGCGGCTAGAGGCTCGGACCACTGACTGGACACCTGCGGGC
 AGCACTGGCCAGGTGGTCCATGGTAGTCCCGCTGAGGGTTTCTGGTGCCTCAACAGGGAGCAGCGGCCCTGGCCA
 GAAGTCTCTAATTACACCGTACGCTTCCCTGCCCCACAGGATCCCTGCGCCGAGACACAGAGCGCATCTGGA
 GCCCATGGTCTCCCTGGAGCAAGTGTCTCAGCTGCCCTGTGGTCAGACTGGGGTCCAGACTCGCACACGCATTGCT
 TTGGCAGAGATGGTGTGCTGTGCAAGTGAAGGCCAGCGAAGAGGGTCAGCACTGCATGGGCCAGGACTGTACAGC
 CTGTGACCTGACCTGCCAATGGCCAGGTGAATGCTGACTGTGATGCCTGCATGTGCCAGGACTTCATGCTTC
 ATGGGGCTGTCTCCCTTCCCGGAGGTGCCCCAGCCTCAGGGGCTGCTATCTACCTCCTGACCAAGACGCTAAAG
 CTGCTGACCCAGACAGACAGTGTATGGGAGATTCCGAATCCCTGGCTTGTGCCCTGATGGCAAAGCATCCTGAA
 GATCACAAAGGTCAAGTTGCCCTTGTACTCACAATGCCAAGACTAGCCTGAAGGCAGCCACCATTCAAGG
 CAGAGTTTGTGAGGGCAGAGACTCCATACATGGTGTATGAACCCCTGAGACAAAAGCACGGAGAGCTGGGACAGC
 GTGTCTCTGTGCTGTAAGGCCACAGGGAAGCCAGGCCAGACAAGTATTTTGGTATCATAATGACACATTGCT
 GGATCCTTCCCTCTACAAGCATGAGAGCAAGCTGGTGTGAGGAACTGCAGCAGCACCAGGCTGGGAGTACT
 TTTGCAAGGCCCAGAGTGTGCTGGGGCTGTGAAGTCCAAGGTGCCCCAGCTGATTGTGTCACAGCATCTGAG
 ACTCCTTGCAACCCAGTTCCCTGAGAGCTATCTTATCCGGCTGCCCATGATTGCTTTTCAAGATGCCACCACTC
 CTTCTACTATGACGTGGGACGCTGCCCTGTAAAGACTTGTGACGGGCAGCAGGATAATGGGATCAGGTGCCGTG
 ATGCTGTGCAGAACTGCTGTGGCATCTCCAAGACAGAGGAAAGGGAGATCCAGTGCAGTGCAGTACACGCTACCC
 ACCAAGGTGGCCAAGGAGTGCAGCTGCCAGCGGTGTACGGAAGCTCGGAGCATCGTGGGGGGCCGTGTGAGTGC
 TGCTGACAATGGGGAGCCCATGCGCTTGGCCATGTGTACATGGGGAACAGCCGTGTAAAGCATGACTGGCTACA
 AGGGCACTTTACCCCTCCATGTCCCCAGGACACTGAGAGGCTGGTGTGCTCACATTTGTGGACAGGCTGCAGAAG
 TTTGTCAACACCACCAAGTGTACCTTTCAACAAGAAGGGGAGTGCCGTGTCCATGAAATCAAGATGCTTCG
 TCGGAAGAGCCCATCACTTTGGAAGCCATGGAGCAACATCACTCCCTGGGGAAAGTGGTTGGTGAAGACC
 CCATGGCTGAAGTGGAGATTCCATCCAGGAGTTTCTACAGGCAGAAATGGGGAGCCCTACATAGGAAAGTGAAG
 GCCAGTGTGACCTTCCCTGGATCCCCGGAATATTTCCACAGCCACAGCTGCCAGACTGACCTGAAGTTTATCAA
 TGACGAAGGAGACACTTTCCCTTCCGACGATGCGATGTTCTCTGTGACTTCAGAGATGAGGTCACTCAG
 AGCCACTTAATGCTGGCAAGTGAAGGTCCACCTTGACTGACCCAGGTCAAGATGCCAGGCACATATCCACA
 GTGAAACTCTGGTCACTCAATCCAGACACAGGGCTGTGGGAGGAGGAAGGTGATTTCAAATTTGAAATCAAAG
 GAGGAACAAAAGAGACAGAACCTTCCCTGGTGGGCAACCTGGAGATTCTGTGAGAGGAGGCTTTAACTTGG
 ATGTTCCCTGAAAGCAGGCGGTGCTTTGTTAAGGTGAGGGCTACCGGAGTGAGAGGTTCTTGCTGAGCAG
 ATCCAGGGGTTGTGATCTCCGTGATTAACCTGGAGGCTAGAAGTGGCTTCTTGTCCAACCTAGGGGCTGGG
 CCGCTTTGACAGTGTGATCAGAGGCCCAACGGGCGCTGTGTCCTGCCCTCTGTGATGACCACTCCCTGATG
 CCTACTCTGCCATGTCTTGGCAAGCCTGGCTGGGGAGGAAGTGAAGCAGTGGAGTCTTCTCCTAAATTAAC
 CCAATGCAATTTGGCGTCCCTCAGCCCTATCTCAACAAGCTCAACTACCGTCCGACGGACCATGAGGATCCACG
 GGTTAAAAAGACAGCTTTCCAGATTAGCATGGCCAAAGCCAACTCAGCTGAGGAGAGCAATGGGCCCA
 TCTATGCCCTTTGAGAACCTCCGGGCATGTGAAGAGGCACCCAGTGCAGCCCACTTCCGGTTCTACCAGATT
 GAGGGGGATCGATATGACTACAACACAGTCCCTTCAACGAAGATGACCTATGAGCTGGACTGAAGACATCT
 GGCATGGTGGCCAAAGCCGATGGAATTCAGGGCCTGCTATATCAAGGTGAAGATTGTGGGGCCATGGAAGTGA
 ATGTGCGATCCCGCAACATGGGGGGCACTCATCGGCGGACAGTGGGGAGGCTGTATGGAATCCGAGATGTGAGG
 AGCACTCGGGACAGGGACAGCCCAATGTCTCAGCTGCCGTGTCTGGAGTTCAAGTGCAGTGGGATGCTCTATGA
 TCAGGACCGTGTGGACCGCACCTGGTGAAGGTCACTCCCCAGGGCAGCTGCCGTGAGCCAGTGTGAACCCCA
 TGCTGCATGAGTACCTGGTCAACCACTTGGCACTTGCAGTCAACAACGACACCAGTGAAGTACACCATGCTGGCA
 CCCTTGGACCCACTGGGCCACAACATATGGCATCTACACTGTCAGTACAGGACCCCTCGCACGGCCAAAGGAGAT
 CGCGCTCGGCCGGTGTCTTTGATGGCACATCCGATGGCTCCTCCAGAATCATGAAGAGCAATGTGGGAGTAGCCC
 TCACCTTCAACTGTGTAGAGAGGCAAGTAGGCCGCCAGAGTGCCTTCCAGTACCTCCAAAGCACCCAGCCAG
 TCCCTGCTGCAGGCACTGTCCAAGGAAGAGTGCCTCGAGGAGGCAGCAGCGAGCGAGCAGGGGTGGCCAGCG
 CCAGGGTGGAGTGGTGGCTCTCTGAGATTTCTAGAGTTGCTCAACAGCCCTGATCAACT**TA**AGTTTGTGGT
 ACTTCACCTCTTCTGCCCTCATTTTCACTGTGACAGCCATTGTGAGACTGATGCACAACTGTCACTTGGTTAAT
 TTAAGCACTCTGTTTTCTGTAATTTGCTTGTGTTTCTTCTTCACTGCCTTTACTTACTTTTGTCCCATGCTACTGA
 TTGGCACGTGGCCCCACAATGGCACAAATAAAGCCCTTTGTGAAACTGTTCTTTAAATGAAACACAAGAAAT
 GGCCACTGGTAAACTCTGCAGCTTCAACTGTACTTCATTTAATGCCATTAATGCAAAATATACTTCTCTCTT
 TTTGATGTTTTTGGCCACCTCTGCAATAGTGATAATCTGATGCTGAAGATCAAATAACCAATATAAAGCATAT
 TTCTTGGCCTTGCTCCACAGGACATAGGCAAGCCTTGATCATAGTTTATACATATAAATGGTGGTGAATAAAG
 AAATAAAACACAATACTTTTACTTGAATGTAAATAACTTATTTATTTCTTTGCTAAATTTGGAATTTAGTGC
 ACATTCAAAGTTAAGCTATTAATATAGGGTGATCATAGTTCCCTCTACCAAGTCTGGAAAGAACATCTCCTGGT
 ATCCACAATTACACCAGGTTGCTAACTGTATTTGTACATTTCCCTTTGCATTGCTTTTGTCTGTAGAAAC
 CCAGTGTAGCCAGGGCAGATGTCAATAAATGCATACTCTGTATTTGAAAAA

FIGURE 72

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNI
DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ
RPGQNCNSNYTVRFLCPPGSLRRDTERIWS PWS PWSKCSAACGQTGVQTRTRICLAEMVSLCS
EASEEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAYLLTK
TPKLLTQTDS DGRFRIPGLCPDGKSILKITKVKFAPIVLTM PKTSLKAATIKAEFVRAETPY
MVMNPETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQQHQAG
EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPE SYLIRLPHDCFQ NATNSFY YDVGRCPV
KTCAGQQDNGIRCRDAVQNC CGISKTEEREIQCSGYTLPTKVAKESCQRC TETR SIVRGRV
SAADNGEPMRF GHVYMGNSRV SMTGYKGTFTLHVPQDTERLVLT FVDRLQKFVNTTKVLPFN
KKGSAVFHEIKMLRRKEPIT LEAMETNIIPLGEVVGEDPMAELEIPSR SFYRQNGEPYIGKV
KASVTFLDPRNISTATAAQTDLNF INDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKV KVLHL
DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL
DVPESRRCFVKVRAYR SERFLPSEQIQGVVISVINLEPRTGF LSNPRAWGRFDSVITGPNGA
CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTDHEDPR
VKKTAFQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFN
EDDPMSWTEDYLAWWP KPMEFRACYIKVKIVGPLEVNVR SRNMGGTHRRTVGKLYGIRDVRS
TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTL VKVIPQGSCRRASVNPMLHEYLVNHLPLAV
NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT
FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVA
QQPLIN

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCTCTCAATATACCTGAATACGCAC
AATATCTTAACCTCTTCATATTTGGTTTTGGGATCTGCTTTGAGGTCCCATCTTCATTTAAAAAAAATACAGAG
ACCTACCTACCCGTAACGCATACATACATATGTGTATATATATGTAAACTAGACAAAGATCGCAGATCATAAAGC
AAGCTCTGCTTTAGTTTCCAAGAAGATTACAAAGAAATTTAGAGATGATTTTGTCAAGATCCCTGTGCATTTTCATG
CCCTTTGGGTTACGGTGTCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTTGTGTAAAGAT
CAGATTTTACACGGAAGAAGGGAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATA
TCTGAAAGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATC
CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCCTGAGCTGATGTTTGATTTT
GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTTGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT
CACTCTGTCTTTGGAGACAAAACCTATTGAGCTAACGACACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC
AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA
GATGCTTTTACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTGCACAGA
AGAGTACTCAACAGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTTCGCGCTTTTGTG
CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAACTCAGAGATTTCTTT
ACAGTCAACAGACCTGAGGATAAGGCTGTTAAGACAGCCGTTGGGGAATATTGTAGATGAGTACACTACCTTTGGC
ACGCTACTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGTGCAAGTGTAACTTCATGCCACTGTATGTG
TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACACTACAGGTCCAGACTGTGGGAAATGCAAGAAG
AATTTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAAATACCTGTATCCC
CAGTATTTCCAGTATTTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA
ACAACGTGCGCTGCCTGTGCCCCGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGCTGGC
AGCTGCGGCTCCGACTCTGGCCAGGGCGCGCCCCCGCACGGCACCCACGGCTGTGCTGTGACTGACCAAGCTGTCT
GGGAACCCGCCAGCCCCCTGGTGTCTAGGCTGTCTCCAGCTCCAGCACACGGCAGGGCTGTGCGCTGGGGGAAGCA
GACACAAACCCAAACATTTGCTACTAACAATAGGAAACACACACATACAGACACCCCCACTCAGACAGTGTACAAA
CTAAGAAGGCCTAACTGAACTAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTCT
TGACTCCAGAGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA
TTGGAAGGCTGGACAGCCCCCCCCAAACAGGAAGAGCAAAAAACAAACAACTCAACCGACCTAAAAACATTTGGC
TACTCTAGCTGGTGGCGCCCTAGTACGATCCGCCCAGTGTGTGGACCAACCAATAGCAATTTTGTGTGTCA
GTGCATTGTGGGCATAAGGAAATCTGTTACAAAGCTGCCATATTGGCTGCTTCCGTCCCTGAATCCCTTCCAAC
CTGTGCTTTAGTGAACGTTGCTCTGTAACCCCTCGTTGGTTGAAAGATTTCTTTGTCTGATGTTAGTGATGCACA
TGTGTAACAGCCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGA
GCACACACCCACTATACAAGAGTGGCTATAGGAAAAAAGAAAGTGTATCTATCCTTTTGTATTCAAATGAAGTT
ATTTTTCTTGAACCTACTGTAATATGTAGATTTTGTATTATTGCCAATTTGTGTTACAGACAATCTGTTAAT
GTATCTAATTCGAATGACCAAGACTGACATTTATTTTGTCTCTTTGCTTGTGTTTGTGTTTCTACTGTGCAGA
GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATAAAGAATATCAGTTTACATATATAACAAGTGAATAAGA
TTCCACCAAAGGACATTCTAAATGTTTTCTTGTGCTTTAACACTGGAAGATTTAAGAATAAAAACTCCTGCA
TAAACGATTTTACGGAATTTGTATTGCAATTTCTTAAGATGAAAGGAACAGCCACCAAGCAGTTTTCACACTCACT
TTACTGATTTCTGTGTGGACTGAGTACATTCAGCTGACGAATTTAGTTCCAGGAAGATGGATTGATGTTCACT
AGCTTGGGACAACTTCTGCAAAATATGAGACTATTTCCACTTGGGAAAAATTACAACAGCAAAAAAAAAAAAAA
AAAAAA

FIGURE 74

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK
VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK
EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMIKESLDYGRTWQPYQYYATDCLDAF
HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEIKDRFALFAGPRLRNMASLYGQLD
TTKKLRDFFTVDLRIRLLRPVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN
SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPKPGTANTCIPSISSIGTNVCDNELLH
CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS
PLVF

TQSTT: 434660

FIGURE 75

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCCGGCTAAGATTGCTGAGGAGGCGG
 CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCGTCCGGGCGAGGTGTCCTCATGACTT
 CTCTTGTGGACC**ATG**TCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC
 CCTCTCAGCCTCTACTGATTTTTACCACACCCAAGATTTTTTGAATGGAGGAGACGGCTCA
 AGAGTTTAGCCTTGCGACTGGCCCAGTATCCAGGTGAGGTTCTGCAGAAGGTGTGACTTT
 AGTATACATTTTTCTTCTTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC
 AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA
 CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTAG
 AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGAATAAT
 TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTCTCACTCTGGAGGACACAGATGTGGCAA
 ATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA
 CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT
 CATTGAGGAGTTACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTCTGCTGGT
 TGGACCAAACCTCG**TGAG**CCAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT
 CCGGGAGCAGTGATGTCAAACCTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA
 AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTCGGAAACATCTGCCCATGTGTATTG
 ATGGCAGAGCTGTTGCCACAAAGCGCTTTTATTTAGGGTAAATAACAAATCCATTCTAT
 TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT
 GCTTCAGAAGTGTTATTTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT
 TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAACCAT
 TTACTTTACATATTCGTTTTCAATACTTGCTGTTTCATGTTACACAAGCTTCTTACGGTTTTT
 TTGTAACAATAAATATTTTGAGTAAATAATGGGTACATTTTAACAACTCAGTAGTACAACC
 TAAACTTGATAAAAGTGTAATAAATGTATAGCCATTTATATCCTATGTATAAATTAAATG
 AGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTTTATTAAAAAAAAAAAAAAAAAAAA
 AAAAG

FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHF
SSFQDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKVW
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

77/330

FIGURE 77

TGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCNTATGACACTACCTGCATTGGCNT
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTGAGAAAGTGAAGTGGCATT
TTAACTATGTAAGTTCCTNTCAGATGGAGTGCAGCTTGGAAAAAATTCAGGAGGAGCTCAAG
TTGCAGCCTCCAGCGGTTCTCANTATGGAGGACACAGATGTGGCAAATGGGGT

77/330

FIGURE 79

MEAPDYEVL SVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDD EDATV NK
IALELCTFTLAIALGAVLLL PFSIISNEVLLSLPRNYIIQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGV LGRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRLLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVILEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMP RGMQGTSLGQVSFSKLGSGGAVIQVVLIFYLMVSSVVG FY
SSPLFRSLRPRWHD TMTQIIGNCVCLLVLSALPVFSRTLGLTRFDLLGDFGRFNWLG NFY
IVFLYNAAFAGLTTLC LVKTFTA AVRAELIRAFGLDRLPLPVSGFPQASRKTQH Q

FIGURE 80

GGCTGCCGAGGGAAGGCCCTTGGGTTGGTCTTGGTTGCTTGGCGGCGGCGGNTTCNTCCCC
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG

FIGURE 81

GACCGACCTTAAAGAGTGGGAGCAAAGGGAGGACAGAGCCTTTTAAAACGAGGCGGTGGTGC
CTGCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTTTC
TGTCGCAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGTGCTTGGCGGCGGCGGCTTCCT
CCCCGTTGTCNTCCCCGGGCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGA
AGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGT
GTATTATATCAACACTTCTGTTTGCAACACTGTACATCNTCTGCCACATCTTCCTGACCCGC
TTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGATTGC
GCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCTCTGCTCCTGCCCTTCT
CCATCATCAGCAATGAGGTGCTGCACTCCC

FIGURE 82

GATGTGCTCCTTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTTT
GGAATTGAGGAACTTCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTTCATGCTGCTGT
GGGTGATATTACTGGTCCTGGCTCCTGTCAGTGGACAGTTTGCAAGGACACCCAGGCCCAT
ATTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGAGTGACCCTCACTTGCAA
GGGATTTGCTTCTACTCACCACAGAAAACAAAATGGTACCATCGGTACCTTGGGAAAGAAA
TACTAAGAGAAACCCAGACAATATCCTTGAGGTTGAGGAATCTGGAGAGTACAGATGCCAG
GCCAGGGCTCCCCTCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTCC
TCATGCTGCCCAGGCTAATGTTGAACTCCTGGGCTCAAGTGATCTGCTCACCTTAGGCCTCTC
AAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGAC
TCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACAAGAA
TGATAATGTCCTGGCATTCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAA
AAA

83/330

FIGURE 83

MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQKTKWYHRYL
GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLDFSSEMGPAAQANVELLGSSDLLT

10571 535660

FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
AQEKFQDLGAAYEVLSDSEKRRKQYDTYGEGLKDGHQSSHGDI FSHFFGDFGFMFGGT PRQQ
DRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGM EYPFFIGEGEPHVDGEPGDLRFRIKVVKH
PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD
NNNIKGLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGGCGGAGGAGGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAGGATATTAAAAAGG
CCTATAGGAACTAGCCCTGCAGNTTTATCCCGACCGGAACCCTGATGATCCACAAGCCCAG
GAGAAATTCCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

FIGURE 87

GGCACGAGGCGGCGGGGAGTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGT
CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA
GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATGG**
AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC
TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATT
TAAGCCCATTTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC
TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC
GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC
AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA
GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG
TACCCTCCGTTGGACCCCCAACTCCTGGACGCACGGACGACTGCCCTGCTCCTGTCTGTGAG
TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC
AGTCTCTGTGCGCTGCTGAGGAGCATTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG
CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATTT**TAGT**GCCT
ACAGGCCAGCAGCTAGCCATGAAGGCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT
CTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTTCAGCTGTGTGTGCATAG
TAAAGCAGGAGATCCCCGTGAGTTTATGCCTCTTTTGCAGTTGCAAACGTGGCTGGTGAGT
GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACCTCTCTGCAAGAGGAGTATTGAAAA
CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTCAGAAAATTGAGCCACCGTCT
AAGAAATCAAGAGGTTTACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG
TGTGGCAATTCTGATCTGCATTTTCAAGAGGACAATCAATTGAAACTAAGTAGGGGTTTC
TTCTTTTGGCAAGACTTGTAATCTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCT
GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGACGGTTTGGGTTTGAAGCTGAGGAACT
ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG
GATAGTAAATTTATACTTATGTTTCCCTCAAAAAAAAAAAAAA

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTCCGAGGTGCTTTCGCCGCTGTCC
CCACCACTGCAGCC**ATG**ATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA
TTTGGAGTGTTTTTCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT
TGGAAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAATTGGTTTAGAAAGAACATTCAGAT
TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTC
CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTGAG
GGGCTTCTTCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCCTGGATCCCTCCTAAAT
TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTA**TAA**CAACA
AGTGAATTTGAAGACTCATTTAAAAATATTGTGTTATTTATAAAGTCATTTGAAGAATATTCA
GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG
CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA
ACTAAGAAGAAGTCAGCAAGCAAAGTGAAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA
ACTCTTGAAGGCTATTTGTGTTGTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA
CTGTGGTGCCTGTTTCTTTTCTTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT
TTTTAGAAGTGTCCTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA
TGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG
ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

FIGURE 90

MISLTDTKIGMGLTGFGVFLLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK
HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFI RRVPVLGSLNLPGI
RSFVDKVGESNNMV

Important features:

Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

1051112233445566778899

FIGURE 91

GAAGACGTGGCGGCTCTCGCCTGGGCTGTTTCCCGGCTTCATTTCTCCCGACTCAGCTTCCC
ACCNTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGCCATGATCTCCTTAA
CGGACACGCAGAAAATTGGAATGGGATTAACCGGATTTGGAGTGTTTTTCCTGTTCTTTGGA
ATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTT
GGCTTTTGTAATTGGTTTAGAAAGAACATTCAGATTCTTCTTCCAAAAACATAAAATGAAAG
CTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATG
ATCTTCGAAATTTATGGATTTTTTCTCTTGTTT

5852555

FIGURE 94

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCG
 GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA
 GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA
 TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG
 GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT
 TGGTACAAATTCCGTCGTCATT CAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAA
 AAAC TTGCAGGAGCATTTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA
 TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT
 TTTTACAGGGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAAC TGGGTTATAAAAC
 TGTATCAGGTTCCGTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT
 TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA
 CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT
 AGTAAAGGATGTAAACAGATTA AAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG
 CAGCAAGAGAGAAGAACATCCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA
 CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTATGTCTTTAAAAAATAGACA
 TGTTTCTAAAAGTAGCTGTA ACTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAA
 TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT
 AAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGA
 CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC
 CAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA
 TTT**TGA**TCCTTTTAACTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAACAT
 TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTAC
 CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC
 ATCAGATGCTTTTATTTCCAAACCTTTTTTTTACCTTTCACTAAGTTGTTGAGGGGAAGGCT
 TACACAGACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA
 TCCCAGCACTTAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCC
 TGGGCAACGTATTGAGACCATGTCTATTA AAAAATAAAATGGAAAAGCAAGAATAGCCTTAT
 TTTCAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG
 TGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA
 ATAAATTTGCAAAACATCATCTAAATTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDRWQFKRSRLLDTOQDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 97

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRRLYVSKGFNKNDAPLHEINGDHLKICPQ
GSTCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF
VKTYGHLYMQNSELFKDLELVELKRYVVGNVNLEEMLNDFWARLLERMFRLVNSQYHFTDEY
LECVSKYTEQLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIES
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGC GPPKPLPAGRISRSISESAFSARFRPHHPPEE
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLF
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGE
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

FIGURE 98

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCT
GACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAA
GCAACTTACAGCTGCACCGACAGTTGCG**ATG**AAAGTTCTAATCTCTTCCCTCCTCCTGTTGC
TGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTGCCAGAGGCCAC
AGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAA
AGATTGGTTCCTGAGAGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGC
AGTCCCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGA
AAGCCAAACAAGCATTCCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTT
TGCTCTGCCTTTG**TAG**GAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAA
GACAGTGAGCACACCTACCAGACACTCTTCTTCTCCCACTCACTCTCCCACTGTACCCACC
CCTAAATCATTCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTC
TCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTT
AATTACCTGAAAGATTCCAGGAACTGTAGCTTCCTAGCTAGTGTCAATTTAACCTTAAATGC
AATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTTAAATGTCAAAAAAAAAAAAAAAAAA

99/330

FIGURE 99

MKVLISSLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR
KFMTVSGLPKKQCPDHFKNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

protein sequence

FIGURE 100

A**ATG**GCTGTCTTAGTACTTCGCCTGACAGTTGTCCTGGGACTGCTTGTCTTATTCCTGACCT
GCTATGCAGACGACAAACCAGACAAGCCAGACGACAAGCCAGACGACTCGGGCAAAGACCCA
AAGCCAGACTTCCCCAAATTCCTAAGCCTCCTGGGCACAGAGATCATTGAGAATGCAGTCGA
GTTATCCTCCGCTCCATGTCCAGGAGCACAGGATTTATGGAATTTGATGATAATGAAGGAA
AACATTTCATCAAAG**TGA**CATCCTCAGGACACACCCATGTGGCTCCTGGACAATCCAAGAGCA
GCCAAATCCTGCTTTTCCAGTTTGGCTCCACAAGTCCTCCAGGACAGAGCCCTCAAAGCAAC
TCCCAACGAGTTCTCAGGATTCAGGCTCTGGCTTCAACCAAACAGAACTCATTTTGAACACC
CTGACTGCATTTTGGCTTTTAGAAAAGTTAGAATAAATATGGCGCTTTGGGATCACATAGTTG
ATGGAGAGGAAA

101/330

FIGURE 101

MAVLVLRRLTVVVLGLLVLF LTCYADDKPDKPDDKPDDSGKDPKPDFPKFLSLLGTEI IENAVE
FILRSMRSTGFM EFDDNEGKHSSK

for seed

FIGURE 102

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT
CAGAGCTGGTCTGCC**ATG**GACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTAC
CCTGCCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAGCTACTTCC
CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG
CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG
CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC
CAAATCCCCACTTTGAGAAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT
GAGCGGTTTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGT
GGTGGTCTGCACTCTGGTGCTGTGCTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGGTCC
GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA
AGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG
CTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG
AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC
AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA
AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCT**TAG**CAGAATGAGAGAAGACATT
CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC
CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC
TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC
CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT
CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT
CTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT
GCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGA
AACCACG

FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAAAGTTGCCTCATCGCAGG
CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA
ACTAATATTTATATGACAGAAGAAAAAG**ATG**TCATTCCGTAAAGTAAACATCATCATCTTGG
TCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG
TTAAGGAATGAGGTTACAGATTGAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAA
TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCCTGTGGTCATCGCTGCATCTG
AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTGAGCACAACACTCGCTCCAAT
GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCTGGCTCAACAG
TGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTTGACCCTAAACTTTTGGAAGGAA
AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC
TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA
AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG
AAGATTGTGATTGAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC
ATTGGCTATCTTGACTATAAAAAGGAAAGAATTCGTAAGCTTTCCATGAAAGCCAGCACTTG
CTCATTTAATCCTGGAGTTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA
ACCAACTGGAAAAATGGATGAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCTGGCT
GGTAGCATCACAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC
TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTA
AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT
ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTCACCTAATCCGAAG
ATATACCGAGATCTCAAACATAAAG**TGA**AACAGAATTTGAACTGTAAGCAAGCATTTCTCAG
GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA
GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGTC
AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTCCTT
ACTACAATGCTGAATGACTGGAAAGAAGAACTGATATGGCTAGTTCAGCTAGCTGGTACAGA
TAATTCAAACTGCTGTTGGTTTTAATTTTGTAACCTGTGGCCTGATCTGTAAATAAACTT
ACATTTTTC

FIGURE 105

MSFRKVNI I I I L V L A V A L F L L V L H H N F L S L S S L L R N E V T D S G I V G P Q P I D F V P N A L R H A V D G R
Q E E I P V V I A A S E D R L G G A I A A I N S I Q H N T R S N V I F Y I V T L N N T A D H L R S W L N S D S L K S I R Y K
I V N F D P K L L E G K V K E D P D Q G E S M K P L T F A R F Y L P I L V P S A K K A I Y M D D D V I V Q G D I L A L Y N T
A L K P G H A A A F S E D C D S A S T K V V I R G A G N Q Y N Y I G Y L D Y K K E R I R K L S M K A S T C S F N P G V F V A
N L T E W K R Q N I T N Q L E K W M K L N V E E G L Y S R T L A G S I T T P L L I V F Y Q Q H S T I D P M W N V R H L G S
S A G K R Y S P Q F V K A A K L L H W N G H L K P W G R T A S Y T D V W E K W Y I P D P T G K F N L I R R Y T E I S N I K

FIGURE 106

TGGTTTTTGCCCCATAAATTCCTCAGCTTGAGCAGTTTGTTAAGGAATGAGGTTACAGATT
CAGGAATTNTAGGNCCTCAACCTNTAGANTTTGTCCCAAATGTTCTCCGACATGCAGTAGAT
GGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATNTGAAGACAGGCTTGGGGGGGCCAT
TGCAGCTATAAACAGCATTTCAGCACAACTCGNTCCAATGTGATTTTCTACATTGTTACTC
TCAACAATACAGCAGACCATNTCCGGTCCTGGNTCAACAGTGATTCCCTGAAAAGCATCAGA
TACAAAATTGTCAATTTTGACCCTAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCA
GGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCCCAGCG
CAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTAC
AATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTAC
TAAAGTTGTCATCCGTGGAGCAGGAAA

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG
GAGCTAGAGGGCAAGTGTGCTCGGCCAGCGTGCAGGGAACGCGGGCGGCCAGACAACGGGC
TGGGCTCCGGGGCCTGCGGCGCGGGCGCTGAGCTGGCAGGGCGGGTCGGGGCGCGGGCTGCA
TCCGCATCTCCTCCATCGCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT
TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT
TTCCGATCACTTCAAAGTGGTTAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT
CCGTAAAGTAAACATCATCATCTTGGTCCTGGGCTGTTGCTCTCTTCTTACTGGTTTTGCAC
CATAACTTCCTCAGCTTGAGGCAGTTTGTAAAGGAATGAGGTTACAGATTCAGGAATTGTAG
GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA
GGAGATTCCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA
ACAGCATTCAGCACAACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA
GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCTGAAAAGCATCAGATACAAAATTG
TCAATTTTGACCCTAACTTTTGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC
ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCACAGCGCAAAGAAGG
CCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA
CTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTACTAAAGTTGT
CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA
TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAAC
CTGACGGAATGGAAACGACAGAATATAACTAACCAACTGGAAAAATGGATGAACTCAATGT
AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAAACACCTCCTCTGCTTATCG
TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT
GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA
TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA
GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA
CAGAATTTGAACTGTAAGCAAGCATTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGAAG
TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG
GTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCAGACAGACTATAGACTATAAAT
ATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA
CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAAACCTGCTGTTGGTTTTAATTTT
GTAACCTGTGGCCTGATCTGTAAATAAACTTACATTTTTCAATAGGTAAAAA

FIGURE 108

CTGCAGGTAGACATCTCCACTGCCCAGGAATCACTGAGCGTGCAGACAGCACAGCCTCCTCT
GAAGGCCGGCCATACCAGAGTCCTGCCTCGGCATGGGCCTCACCATTGAGGCAGCTCCACTG
TCTGTGCTGGTCTGAGGGTGCTGCCTGT**ATG**GGGGCAGCCATCTCCCAGGGGGCCCTCATC
GCCATCGTCTGCAACGGTCTCGTGGGCTTCTTGCTGCTGCTGCTCTGGGTCATCCTCTGCTG
GGCCTGCCATTCTCGTCTGCCGACGTTGACTCTCTCTCTGAATCCAGTCCCAACTCCAGCCC
TGGCCCCTGTCCTGAGAAGGCCCCACCACCCCAGAAGCCCAGCCATGAAGGCAGCTACCTGC
TGCAGCCCTGAAGGCCCTGGCCTAGCCTGGAGCCCAGGACC**TAA**GTCCACCTCACCTAGAG
CCTGGAATTAGGATCCCAGAGTTCAGCCAGCCTGGGGTCCAGAACTCAAGAGTCCGCCTGCT
TGGAGCTGGACCCAGCGGCCCAGAGTCTAGCCAGCTTGGCTCCAATAGGAGCTCAGTGGCCC
TAAGGAGATGGGCCTGGGGTGGGGGCTTATGAGTTGGTGCTAGAGCCAGGGCCATCTGGACT
ATGCTCCATCCCAAGGGCCAAGGGTCAGGGGCCGGGTCCACTCTTTCCCTAGGCTGAGCACC
TCTAGGCCCTCTAGGTGGGGAAGCAAACCTGGAACCCATGGCAATAATAGGAGGGTGTCCAG
GCTGGGCCCCCTCCCCTGGTCCTCCCAGTGTTTGCTGGATAATAAATGGAACCTATGGCTCTAA
AAAAAAAAAAAAAAAAAAAA

109/330

FIGURE 109

MGA AISQGALIAIVCNGLVGFLLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRPHH
PRSPAMKAATCCSPEGPWPSLEPRT

109/330

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

111/330

[illegible]

FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVTYPSALTYSSSKSPAAQAGETTKAYQRPPPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTTSSQNRPRADPGIQ
RQDPGGAFFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHNTSRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTLLQPLVKRVCDTDRLACSKTCLNSADI
GFVIDGSSSVGTGNFRTVLQFVTNLTKFEFISDTRIGAVQYTYEQRLEFGFDKYSSKPDIL
LNAIKRVGYWSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

FIGURE 114

CAGGATGAAC TGGTTGCAGTGGCTGCTGCTGCTGCGGGGGCGCTGAGAGGACACGAGCTCTA**TAA**
TGCCTTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT
GCGCCAGGTCCCGACGGCTCCGCGCCAGATCCCGCCCCTACAGTTTTTCTCTGACTCTAAT
TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATTCCAAAGAGTG GTTGAAG
TGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCTGTGTTTGAAACAAACATT
CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA
GGCTGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAACTCC
TCCCAGCCTTTCAGACCCCCACTGGCATGCCATATGGAACAGTGAACCTTACTTCATGGCGTG
AACCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC
CACCTGAGCAGCCTCACTGGTGACCCGGTGTTTGAAGATGTGGCCAGAGTG GCTTTGATGC
GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTTCGGCAACCACATTGATGTGCTCACTGGC
AAGTGGGTGGCCAGGACGCAGGCATCGGGCTGGCGTGGACTCCTACTTTGAGTACTTG GT
GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCTAGAGTATAACAAAG
CCATCCGGAACCTACACCCGCTTCGATGACTGGTACCTGTGGGTTGAGATGTACAAGGGGACT
GTGTCCATGCCAGTCTTCCAGTCCTTGAGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG
AGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGG
GGCTCCCGGAATTCTACAACATTCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA
CTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCCCT
CCTAGAACTCGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGAT
TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG
GCCGAGACTGTGAAATACCTCTACCTCCTGTTTGACCCAACCAACTTCATCCACAACAATGG
GTCCACCTTCGACGCGGTGATCACCCCTATGGGGAGTG CATCCTGGGGGCTGGGGGGTACA
TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG
GAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCTCAAACGGAGCAGGTC
GAAATTTTCAGAAAAACACTGTTAGTTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC
TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCCA
CTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTTGGCATTACTGGGACAGGTTTTCTCT
AGACTCCTCA**TAA**CCACTGGATAATTTTTTTTATTTTTATTTTTTTGAGGCTAAACTATAATA
AATTGCTTTTTGGCTATCATAAAA

FIGURE 115

MPFRLLIPLGLLCALLPQHGGAPGPDGSAPDPAHYSFSLTLIDALDTLLIILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHLSSKKAGVEVEAGWPCSGPLLMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLVVQMYKGTVSMFVQSLEAYWPGLSLIGDIDNAMRTFLNYYTVWKQFG
GLPEFYNIPOGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKL DNRMESFFLAETVKYLYLLFDPTNFIHNNGSTF DAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

115/330

FIGURE 116

AAAGTTACATTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTCTGAATACCAGGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTC**TAG**GTTTTCGGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGA CTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTCTAG
TGTTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGAGAGAGCCCACTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGGTGACAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAAA

117/330

FIGURE 117

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS
SILKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMILIV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

FIGURE 118

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCCAC
ATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 119

CGGACGCGTGGGCCGCCACCTCCGGAACAAGCC**ATG**GTGGCGGCGACGGTGGCAGCGGCGTG
GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG
TCAACATCCGGGGCAAACCTGGTGTGCTGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG
AATGTGGCCAGCGAGTGCGGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG
AGACCTGGGCCCCCACCCTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG
AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCCGCCGCACCTACAGTGTCTCATTCCCC
ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCCA
GACTTCTGGGAAGGAGCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG
TGGTAGGGGCTTGGGACCCAACTGTGTGCTAGTGAGGAGGTCAGACCCCAGATCACAGCGCTC
GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTA**TAA**CCACCGCGTCTCCTCCTCCACCA
CCTCATCCCGCCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGCTTCAAAGGGAG
AGACCCACTGACTCTCCTTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA
AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACCTCCTGGCCAATGAG
AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA
TAGAAGTATATCAAGCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC
CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC
CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT
AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT
GTTATTTCTCCTCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA
AACAAATACCTCACGATATAAAATAAAAATGAAAGTATCCTCCTCAAAAA

120/330

FIGURE 120

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFTDQ
HYRALQQQLQRDLGPHHFNVLAFFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG
AHPAFKYLAQTSKGKEPTWNFWKYLVPDVGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

120/330

FIGURE 122

MAVYVGMLRLGRLCAGSSGVLGARAALSRWQEARLQGVRFSSREVDRMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLPGTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCCKIVGRSKDMIIRGGENIYPAELEDFHHTHPKVQEVQVVGVKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

Signal Peptide:

amino acids 1-22

Transmembrane Domains:

amino acids 140-161, 213-229, 312-334

Putative AMP-binding Domain Signature:

amino acids 260-271

N-myristoylation Sites:

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277,
314-319, 318-323, 379-384, 380-385, 409-413

N-glycosylation Site:

amino acids 282-285

FIGURE 123

CAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGACACCAGAGCAGTTGCGGA
TGATCCTGCCCCAACCCCCTGTACCATTGCCTGGGTTCCGTGGCAGGCACAATGATGTGTCTG
ATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGCAAGAAGGCACTGGAGGC
CATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGATGTTCTGTGGACATTCTGA
ACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAGGTGTCATTGCTGGGTCC
CCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAATATGAAGGACCTGGTGGT
TGCTTATGGAACCACAGAGAACAGTCCCGTGACATTTCGCGCACTTCCCTGAGGACACTGTGG
AGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGGCGCGGATCATGAACATG
GAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGCATCCGAGGGTACTGCGT
CATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGTGGATCAGGACAAGTGGT
ATTGGACAGGAGATGTCGCCAC

FIGURE 124

GAGCAGGACGGAGCCATGGACCCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGC
AGGCTGGCTGCTGCTGCTGCTGCTTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG
TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC
GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC
AGTGCGGGGTTGCGGTTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGG
TTCTGGCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC
ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTG
CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCTGTGAGCT
GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCA
GCTAATGTGACTGTGTCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA
TGGAGTAACAGGCCCCAGGGTTACGCTCAGTGGCTCCTGTTGCCAGGGGTCCTCGCTGTAAC
CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT
CCAGAGCCACGACTGTGGCCTCAACCACATCTGTCAACCACTTCTACCTCGGCCCCAGTGAG
ACCCACATCCACCACCAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG
AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC
CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG
TGTGGCTCCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCTACTGTT
GAGCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT
CATCACTTCTGTTCCCAACCACTGGACTGGGCTGGCCAGCCCCGTGTTTTTCCAACATTCCC
CAGTATCCCCAGCTTCTGCTGCGCTGGTTTTGCGGCTTTGGGAAATAAAATACCGTTGTATAT
ATTCTGCCAGGGGTGTTCTAGCTTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC
TCCGCTTGTCTCTTGTGATGTTAGGACAGAGTGAGAGAAGTCAGCTGTCACGGGGAAGGTG
AGAGAGAGGATGCTAAGCTTCCCTACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGG
GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG
GAATCGGTTCCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGGCCCGGTAC
CCAATTCGCCCTATAGTGAGTCGTA

125/330

FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCS PNKMKT VKCAPGVDVCT
EAVGAVETIHGQFSLAVRGCGSLPGKNDRGLDLHGLLAFIQLQOCAQDRCNAKLNLT SRAL
DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCFDGNVTLTAANVTV
SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRI PPLVRLPPPEPTT
VASTTSVTTST SAPVRPTSTTKPMPAPTSQT PRQGV EHEASRDEEPRLTGGAAGHQDRSNSG
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

0997535 11501

FIGURE 126

[illegible]

127/330

FIGURE 127

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRRCK
CSFNQKPRAPGDEEAQVENLITANATEPQKQRTEVQPSGGSLWNLRRLLEPLDANVDA

127/330

FIGURE 128

AAACTTGACGCCATGAAAGATCCCGGTCCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT
CCTCTGCCCAGGGAGCCACCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG
CGTCACGACCCGAGGCCTTTAACACCCCGTTCCTGAACATCGACAAATTGCGATCTGCGTTT
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAAGGAACTTCCTTT
CCTCAACTGGGATGCCTTTCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAGT
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT
CTTTCCTGCCTCAGGAACTCCAATAAAACATTTTCCATCCAAA

129/330

FIGURE 129

MKIPVLPVAVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSFAKDE
FLNWHALFESIKRKLPFLNWDAFPKLGKGLRSATPDAQ

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 130

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCC**ATG**GGGGTGGAGATTGC
CTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTC
TTCTCCAGCCAGTTCCTCAACTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGC
TTTGCCGGCCACTCA**TGA**GAGTGTTTTTGTGTAAAGTATTTTTTAGAATACTGTTGACTTCT
TCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCA
AATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTT
AGGTTCAAGGTTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTTCATC
TTCATATTAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTT
TATCCCTGGGTGCCCCTGACACATTTATGTAGTGATCCACAAATGTGATTGTTAATTTAAA
TGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTT
AAAAGTTTTGAGTATATTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

10541534660

131/330

FIGURE 131

MGVEIAFASVILTCLSLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

protein

FIGURE 132

GGGGAATCTGCAGTAGGTCTGCCGCGCATGGAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG
GCTGCTGTTGTTCCCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAATGGA
AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAAC
TGCAGCTGCTACCATGGTGT CATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG
GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA
GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTGAGCACTTTATT
TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA
GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC
ATGATATCATGTATCCTGCTTGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT
CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG
GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG
ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAACTTGTTGATGCAGAATACACCAAAAAC
CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT
GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA
AACACCTCTTCCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC
TATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAAACAGATCTCTCCAATGTCCA
AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA
GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG
AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT
TCCCAAAATGTTGAAAAC TGA ACTA TAG TAGTCATCATAGGACCATAGTCCTCTTTGTGGCA
ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA
TCTGCTATCAAGCCAAATACCTGGTTTTTCCTTATCATGCTGCACCCAGAGCAACTCTTGAGA
AAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTTGGATGAATAAGGA
CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC
ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA
TGTGATGATGCCCTTTGTCCCATTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA
CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG
AAACCCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG
TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTGTAAAACCATAAACTCTGTTACTCAG
GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT
TCAGGTTCCCTTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

FIGURE 134

CACCCCTCCATTTCTCGCCATGGCCCCCTGCACTGCTCCTGATCCCTGCTGCCCTCGCCTCTT
TCATCCTGGCCTTTGGCACCGGAGTGGAGTTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT
GGAGGGATCCCGGAGTCTGGTGGTCCGGATGCCCGCCAGGGATGGCTGGCTGCCCTGCAGGA
CCGCAGCATCCTTGCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC
ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCCTTCAG
AGGTCACCTGTATGTGGCCTGCACTGCCCTGGCCTTGACAGCTGGTGTATGCGGTACTGGGAGCC
CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCG
TCCTCTGCTTTGTGCTCCATGTCACTCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT
GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCGAGCC
TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCACCTGCGCCACCCAGTGTGTG
TGGAGCTGCTGACAGTGTGTGGGTGGTGCCTACCCTGGGCACGGACCGTCTCCTCCTTGCT
TTCCTCCTTACCCTCTACCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT
CCGGGCCCAGCTACAAAGAAAACCTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAGTT
GAGGAGCTCACTCTGGTTACAAGCCCTGTTCTTCTCCTCTCCCACTGAATTCTAAATCCTTAAC
ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCCTT
CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA
CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGTATGGAAGTCATCTGTCCC
TTCCTGTTTTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC
CTGACCACTCCCCTGGCACTGTTACTTGCTCTGCGCCTCAGGGGTCCCCTTCTGCACCGCT
GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA
GGCCCCAACCTTGCCCTCACCACTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT
GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGCTC
CCTGCTCTCCCGAGGAAGATGCTCTGCAGGAAAATAAAAGTCAGCCTTTTTTCTAAAAAAA

FIGURE 135

MAPALLLIPAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAALQDRSILAP
LAWDLGLLLLFVQGHSMAAERVKAWSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV
LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP
RALRLFSLRHPVCVELLTVLWVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR
KLHLLSRPQDGEAE

Signal sequence:

amino acids 1-13

Transmembrane domains:

amino acids 58-76, 99-113, 141-159, 203-222

N-myristoylation sites:

amino acids 37-43, 42-48, 229-235

FIGURE 136

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA
AGAAATTGCCAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGT
TTTAATTAAGCATGGAATACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATT
CCACAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTA
TCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTAAGAAGCCTTCGGCTCGCGTGCTTCTG
AGCTGCTGTGGATGCGCTCGGCTCTCTGGACTGTCCTTCCGAGTAGGATGTCAGTGAATCC
CTCAAATGGAGCCTCCTGCTGCTGTCACCTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT
TCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTT
ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTT
CTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTAC
TTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAG
AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT
GACATAATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGC
ATTCAGGTGGGTAACTGAGTTTTGCCCAATGCCAAGTACGTAATGAAGACAGACACTGATG
TTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT
TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT
TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATATA
TGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAGAACCCATCAAGTTT
GAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAAAGTGAACATTCATATTCCAGAAGA
CACAAATCTTTTCTTTCTATATAGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTG
CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTGGCAGGTCATGCTAAGGAACACC
ACATGCCATTATTAACCTTCACATTCTACAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA
GTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTG
AACTGAAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC
AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA
GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGG
TGTTACTGAGTTATAAGCTCACTAGGCTGTAAAAACAAAACAATGTAGAGTTTTATTTATTG
AACAAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA
TGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT
TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTTCACAGTTATTATTATTTAAATTA
CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT
CATTCCTTTACATGCAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC
TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAAT
ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAA

137/330

FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQD
FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEA EK
EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTD TDVFIN
TG NLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD
LVPRIYEMMGHV KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG
FSSKEIITFWQVMLRNTTCHY

0907585 44504

FIGURE 138

CCTCTGTCCACTGCTTTCGTGAAGACAAGATGAAGTTCACAATTGTCTTTGCTGGACTTCTT
GGAGTCTTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACAACAA
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA
ATAACAACGGATGGGACTCCTGGAATTCCATCTGGGATTATGGAAATGGCTTTGCTGCAACC
AGACTCTTTCAAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCATGCCCTCCAT
TCAATCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGACCAGGAGGACCAC
CTCCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACCTGAGCAAGTTCGGA
AAAAACATTGCAAACATGTGTCGTGGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC
AAGCCTGTTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGTGGACATTT
CCTTCTGTGGAGACACGGTGGAGAACTAAACAATTTTTTAAAGCCACTATGGATTTAGTCAT
CTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTTTTTACCATGTCATTCTGAAATT
TTTCTCTACTAGTTATGTTTGATTTCTTTAAGTTTCAATAAAATCATTTAGCATTGAAAAAAA

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDNNDNAGSGQSVSVNNEHNVANVDNNGWDSWNS
IWDYGNNGFAATRLFQKKTICIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLMYSVN
PNKVDDLKFKGNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

139-140

FIGURE 140

CATTTCTGAACTAATCGTGTGAGAATTGACTTTGAAAAGCATTGCTTTTTTACAGAAGTATA
 TTAACTTTTTAGGAGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCG
 CTCATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTCTTATTGCTTACTGATTTTTT
 TGAGTTAAGAGTTGTTATATGCTAGAAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAA
 GAATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAACTGGTTTGTTTACATG
 CAAGCTTATAGTTGAAATATTTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGT
 TTGTTTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTT
 CAGATTCGGTTGCCAACTCGTCCCCATTGGTTTCTTCTTTTTTGGTACTACAGAAGAGGAAAT
 CCAGGAAATCTGCATAGAAACACTTAGGCTTTATAACCAGAAAAAAGCCAAACTATGAATTAC
 TGGAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCAAATTTAAAAGCAAAGGGA
 TTGAATCCGGATGGAAGTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCCAGCCTCCAAGCC
 ATCATCACCAAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAG
 TCAAAAAAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAGAAAA
 GACAGCAAGAGAAGTAGAAATAGCAGAAGTGCAAGTCGATCGAGGTCAAGAACACGATCACG
 TTCTAGATCACATACTCCAAGAAGACACTATAATAATAGGCGGAGTCGATCTGGAACATACA
 GCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTCGAAGACATCATAATCAT
 GGTTCCTCCTCACCTTAAGGCCAAGCATACCAGAGATGATTTAAAAAGTTCAAACAGACATGG
 TCATAAAAGGAAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAG
 CCAAGAAACACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTT
 GAGAGGTCCCATAAAAGCAAGCACCATGGTGGCAGTCGCTCAGGACATGGCAGGCACAGGCG
CTGACTTTCTCTTCCTTTGAGCCTGCATCAGTTCTTGGTTTTGCCTATCTACAGTGTGATGT
 ATGGACTCAATCAAAAACATTAAACGCAAACCTGATTAGGATTTGATTTCTTGAAACCCTCTA
 GGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTTTGCACATT
 AAAATGCCCTAGCAGTATCTAATTTAAAACCATGGTCAGGTTCAATTGTACTTTATTATAGT
 TGTGTATTGTTTATTGCTATAAGAAGTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTT
 ATACAGATAAAATTGCAGACACTGTTCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAAT
 ACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATTTTTTACAAGGAAATAAAATACAAAT
 CTTGTTTTTTTCTAAAAAAGT

FIGURE 141

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLY
TRKKPNYELLEKEVEKRVKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEK
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYN
NRRSRSGTYSSRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRKKSRSRSQ
SKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSRGHGRHR

FIGURE 142

TGGGGATAAAGGAAAAATGGTCAGGTATTAATGGCTTAAAGATTATTGGAAGGGGTTTATCA
TTTTTTGAANNTATTCGGGTCANAATTGNCTTTGAAAAGCATTGCTTTTTACAGAAATATAT
TANCTTTTTAGAGTAATTTCTAGTTTGGATTGTAATATGAAATTATTTAAAAGGGCTTCGCT
CATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTNNTATTGCTTACTGATTTTTTTTG
AGTTAAGAGTTGTTATATGNTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAGA
ATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGCTGGTTTGTTTACATGCA
AGCTTATAGTTGAAATATTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTT
GTTTCGATTTCAACCAGAGANTATAGCATGTGCTTGCATCTACCTTGCAGNTAGAGCACTTCA
GATTCCGTTGCCAACTNGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAATCC
AGGAAATNTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAACTATGAATTACTG
GAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCNAATTAAAAGCAAAGGGATT
GAATCCGGATGGAACTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCC

FIGURE 143

GGCACGAGGCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGTC**ATGGC**
GGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC
CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTTTGTTCGGATACAAGCACCCG
TCTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCCAGAGGCAGGAAAGAGCGGTG
GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG
AGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG
TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACTACAT
GCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGTGCCTGCTCACGGTGACCTTCT
CCATCAAGATGTTCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC
TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGCCATGCTGGTGCAAGTGGTGCG
GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC
CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG
GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC
CCAGACCCACCGGGACGCACTGACCATGTGCGAGGACAGACCCATGCTGCAGTTCCCTCCTGC
ACACCAGCTTCCTGTCTCCCCTGTTATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC
TTCCTGCACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA
CTCTGGGCGCCTCTGGTTGCTGGTGGTGTGCTGTGCCTGCTGCGGCTGGCGGTGACCCGGCCCC
ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC
CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT
GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCTCAACTGCACACTTCTGCTCAAGACGC
TGGGAGGCTATTCTGGGGCCTGGGCCCAGCTCCTCTACTATCCCCCGACCCATCCTCAGCC
AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG
GGCCCTGGGTGGCCTGCTTACTCCCCCTCTTCCTCCGTGGCGTCTGGCCTACCTCATCTGGT
GGACGGCTGCCTGCCAGCTGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA
GGCTCC**TAG**CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCTGGGGCAGCGGGACA
CTAGCCTGCCCCCTCTGTTTGCGCCCCCGTGTCCCCAGCTGCAAGGTGGGGCCGGACTCCCC
GGCGTTCCCTTCACCACAGTGCTGACCCGCGGCCCCCTTGGACGCCGAGTTTCTGCCTCA
GAACTGTCTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG
TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC
GATTTTAA

FIGURE 144

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEEELRALAGKPRPRGRKE
RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY
YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV
VREETLELGLPEGLASMTQNLEPLLKKQGWDWALPVAKLAIRVGLAVVGSVLGAFLTTFPGLR
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT
VVSLQYLTPILITLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAAARI
AGALGGLLTPLFLRGVLAYLIWWTAAACQLLASLFGLYFHQHLA

FIGURE 145

CGTTNGCACGCGTCAATGGCGGTCTCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTT
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCC
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC
GAGATGCCCCGTTCCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCCTGCGC
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGT
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTCCCTGACAGTGACACGGCTGTACTTCAGC
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGC
CATGCTGGTGCAAGCG

GGTTCCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTTATTAACGTGGCTT
AATCTGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGTTTGTCTTAAAGGAGC
TTGGCTGGTTTTGGGCCCTTGTAGCTGACAGAAAGTGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGG
CGCTTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCTCTG
TATTCAAGAACTCTGTAAAGGTCCTCCCCTACCGGCTGACCAAAAGATAGGAAGAGCGCTCACAAGATGGCTG
TCCAGACGGTCTGCGAGCCTCAGAGCCAGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATTCTCTTAA
TGACAGACGAGCCTGGCCTAGACAACCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCA
GTGGACTCTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAA
AAAAATAAATCGAGCTTTGAGTGTTCTTGAAGGACAAAGACGGGAGTGCAGTTGCCAACCATGCCGACCAGG
GCAGGGAAAAATTCTGAAAAACCACTGCCCCGAAAGTCTTTTCAAGGTTGTACCACCTGATTCCAGATGGTGAA
ATTACCAGCATCAAGATCAATCGATAGATCCCAGTGAAGCCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAAC
CCCATTGGTCCATATTATTATCCAAACATTTATCTGTGATGGGTGATCGCCAGAGACGGCCGGCTACTGCCAG
GAGACATCATTTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACCTACGCTGTGCGTCTCTGCGG
CAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCGACGAGGAACAATGGACAGGCCCC
GGATGCCCTACAGACCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCCCGAGGAGCAGCTTGGAA
TAAACTGTTGTCGCAAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCCATCGA
CATGGTCAGCTTGAAGGAAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCAGAAG
TGCGGCTCATCTGATTCAGGCCAGTGAAGACGTGTTCACTCGTGTCCCGCCAGGTTCCGCGAGCGGAGCC
CTGACATCTTTCAAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACT
CCCAAGCCCCCTCCATCTACAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCT
CGGCATGACCGTCGAGGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTGATCAGTGTTGAGCCG
GAGGAGTCATAAGCAGAGATGGAAGAATAAAAAACAGGTGACATTTTGTGTAATGTGGATGGGGTCGAACCTGACA
GAGGTGAGCCGGAGTGAGGCAGTGGCATTTTGAAGAACAATCATCTCGATGACTCAAGCTTTGGAAGT
CAAAAGATGATGAGCCCCAGGAAGCTGCGACGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCC
GTGACTGGTCCCCATCTGGGTGATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTA
CGAAGAAACACAGCTGGAAGTCTGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAACAACCTTT
TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTG
CTGTCAATGGTAGAAGTACATCAGGAATGATACATGCTTGCTTGGCAAGACTGGTGGAAGAAGCTAAAGGAAGA
ATTACTCTAACTATTGTTTCTTGCCCTGGCATTTTTTTATAGAACTCAATGCGTCAAGAGAAAACAGAAAATA
TCACAAATAGGCTAAGAAAGTTGAAACACTATATTTATCTTGTGAGTTTTTATATTTAAAGAAAGAATACATTGT
AAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAATTA
AAACTACTAGTTTTTTTTTTCAGTGTGGAGGATTTCTATTACTCTACAACATTGTTTATATTTTTTCTATTCAAT
AAAAAGCCCTAAAAACACTAAAAATGATTGATTTGTATACCCCACTGAATTCAAGCTGATTTAAATTTAAATTT
GGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTATTTTATTTTACAGCTAAAATATTTTTTAAATGCA
TTGCTGAGAAAACGTTGCTTTTCATCAAAACAAGAAATAAATATTTTTTCAAGAGTTAAA

FIGURE 148

CCAAAGTGATCATTTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCT
GCACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAAATGGTGCTGACCATCT
TTGGGATACAATCTCATGGATACGAGGTTTTTAACATCATCAGCCCAAGCAACAATGGTGGC
AATGTTTCAGGAGACAGTGACAATTGATAATGAAAAAATACCGCCATCGTTAACATCCATGC
AGGATCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGC
TCTCCCGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAAT
CTCCAATGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAAATACACCTG
GGTCAAGTACAACCCCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTTGGGTCAC
CCATTGAGAACTCTGCAAACATATCCCTTTGTATAAGGGGGAAGTGGTTGAAAACACACAT
AATGTCGGTGCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTG
TGCAGACATTCATGTTTAGGATGATTAGCCCTCTTGTTTTATCTTTTCAAAGAAATACATCC
TTGGTTTACACTCAAAAGTCAAATTAAATTCTTTCCCAATGCCCCAACTAATTTTGAGATTC
AGTCAGAAAATATAAATGCTGTATTTATA

FIGURE 149

MKILVAFLVVLTIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT
IFDYKHGYIASRVLSRRACFILKMDHQNIPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

FIGURE 150

GGCACGAGCCAGGAAC TAGGAGGTTCTCACTGCCCCGAGCAGAGGCCCTACACCCACCGAGGC
ATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG
CCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA
GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCCTGCCG
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA
GCCCCATCCTTGCCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG
GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG**TAG**AATGAACC
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTTCGTATTTGGA
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAA

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVFPKGRWVLITCCAPQPPPPITYSL
CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK
PVSELRANFTLQDRGAGPRVEMICQASSGSPITNSLIGKDQGVHLQQRPCHRQPANFSFLP
SQTSDWFWCQAANNANVQHSALTVPVPPGGDQKMEDWQGPLESPILALPLYRSTRRLSEEEFG
GFRIGNGEVRGRKAAAM

Signal Peptide:

amino acids 1-18

N-glycosylation Sites:

amino acids 86-89, 132-135, 181-184

FIGURE 152

GGTCCTTAATGGCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG
CTGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGT
CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA
CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA
CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT
TACAGAGCAACTGCGTGACATTGAGCTGGAGAATTACACACCCAAGGAACCCCTCACCTGC
AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT
TTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC
TGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATT
ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC
CTGGAGCCAAGTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCAC
AGCCACCACCCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA
TCTGAGGAGAGTCCTTTAGAGTGACAGGTAAAGCTGATACCAAAGGCTCCTGTGAGCACG
GTCTTGATCAAACTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT
GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCCTT
TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC
TGATGGAATTCCTGCACTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTC
TCTTTTTGTTTGGAAAATCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATT
GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG
AATTTTTAAATTATTTAATAAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT
TGTTCTGTACTGATATTTAAATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 153

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFL
HYDCGNKTVTPVSPLGKKLNVTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR
MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS
MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

Important features:**Signal peptide:**

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

155/330

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKKLA
EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGQPTQHFWARL

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT
RRGLSEVPQGI PSNTRYLNL MENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY
ISEGAFEGLFNLKYLN LGMCNIKDMPNLTPLVGLEELEM SGNHFP EIRPGSFHGLSSLKKLW
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLP HDLFTPLRYLVELHLHNPWNCD CDILW
LAWWLREYIPTNSTCCGRCHAPMHRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSVKWLLPNGTVLSHASRHPRISVLNDGT LNF SHVLLSDTGVYTCMV TNVAGNSN
ASAYLNVSTAE LNTSNYSF FTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATDTTDMQTS LDEV MKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRS
TVTAARTVEIIQVDEDI PAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNSLHPTVTTISEPYIIQTHTKDKVQETQI

FIGURE 158

CGCTCGGGCACCAGCCGCGCAAGGATGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCC
 TTCTCATCTCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATG
 TGTCGGGAGTGCTGTGAATATGATCAGATTGAGTGCGTCTGCCCCGAAAAGAGGGAAGTCGTGGGTATACCAT
 CCCTTGCTGCAGGAATGAGGAGAATGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAAC
 GCAAGAGCTGCCGAAATGGCTCATGGGGGGTACCTTGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAG
 TGCCGAGCAGGCTGGTACGAGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTT
 GTTGGAAAGCTATCCCCATAATGCTCACTGTGAATGGACCATTATGCTAAACCTGGGTTTGTCCATCCAATAA
 GATTTGTCTGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAGGTTCTGTGATGGAGACAAC
 CGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAGGATCCTCACT
 CCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGCAT
 GTCCTCATCCCCCTGTTTCCATGACGGCACGTGCGTCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTG
 GCAGGCTATACTGGGCAGCGCTGTGAAAATCTCCTTGAGAAAAGAACTGCTCAGACCTGGGGGGCCAGTCAA
 TGGGTACAGAAAATAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAAATGGCACCGTGGTGTCTT
 TCTTTTGTAAACAACCTCTATGTTCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAAATGGAGAGTGCTCAGG
 AAACAGCCCATGTCATAAAAGCCTGCCGAGAACCAGAAATTTAGACCTGGTGAGAAGGAGAGTTCTTCCGAT
 GCAGGTTCAGTCAAGGGAGACACCATTACACCAGCTATACTCAGCGGCCTTCAGCAAGCAGAACTGCAGAGTG
 CCCCACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCATGGGATACCAACATCTGCATACCCAGCTCCAG
 TATGAGTGCATCTCACCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGTCTGAGGACTGGGAAGTGGAG
 TGGGCGGGCACCATCCTGCATCCCTATCTGCGGGAAAATGAGAACATCACTGCTCCAAAGACCCAAAGGGTGC
 GCTGGCGGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGG
 TTCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTACTGACCTGGG
 GAAGGTACCATGATCAAGACAGCAGACCTGAAAGTTGTTTGGGGAAAATTTACCGGGATGATGACCGGGATG
 AGAAGACCATCCAGAGCCTACAGATTTCTGCTATCATTCTGCATCCCACTATGACCCCATCCTGCTTGATGCT
 GACATCGCCATCCTGAAGCTCCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCTCGCTGCCAG
 TCGGGATCTCAGCACTTCTTCCAGGAGTCCCACATCACTGTGGCTGGCTGGAATGTCTTGGCAGACGTGAGGA
 GCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGGTGGTCACTGTGGTGGACTCGCTGCTGTGTGAGGAGCAG
 CATGAGGACCATGGCATCCCAGTGAGTGTCACTGATAACATGTTCTGTGCCAGCTGGGAACCCACTGCCCTTC
 TGATATCTGCACTGCAGAGACAGGAGCATCGCGGCTGTGTCTTCCCGGGACGAGCATCTCTGAGCCACGCT
 GGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCACTGCCTTACCAAG
 GTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAGTGTTTC
 TGTATATCCGTCTGTACGTGTGTCTTGCCTGAAGCAGTGTGGGCTGAAGTGTGATTTGGCCTGTGAACTTGG
 CTGTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGAT
 GCCGCGTCCACTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAAGAAGACC
 ATATACAAAACCTCTCCACTCCACTGACCTGGTGGTCTTCCCCAACTTTCAGTTATACGAATGCCATCAGCTTG
 ACCAGGGAAGATCTGGGCTTCATGAGGCCCTTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCC
 CAGGGCAGCAGAGCTGGGATGTGGTGCATGCCTTTGTGTACATGGCCACAGTACAGTCTGGTCCTTTTCCTTCC
 CCATCTCTTGTACACATTTTAATAAAATAAGGGTTGGCTTCTGAACTACAAAAAAAAAAAAAAAAAAAAAAAAA
 AA
 AA

FIGURE 159

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVV
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGFYCAECRAGWYGGD
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDG
TCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV
VSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLH
QLYSAAFSKQKLQSAPTCKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE
RTVVVAAHCVTDLGKV'TMIKTADLKVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLD
ADIAILKLLDKARISTRVQPICLAASRDLSTS FQESHITVAGWNVLADVRS PGFKNDTLRSG
VVSVD SLLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR
WHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA
AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAAAC**ATG**GGC
TTCAACCTGACTTTCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT
GACAGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAG
CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACT
AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACCTGTCCTTCTGTGTCTCCTTACCTCAG
AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC
CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC
CTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT
CCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT
TTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC
TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA
GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG
GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCT
AACAACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG
AATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG
ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTACGAGTCTGG
AGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA
TATCAACATCACAGTGGATTTCTGGTTTGGTGCAT**TGA**CCCTGGATCTTTTGGTGATGTTTGG
AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA
AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCTTTTTGTATTTTCTTAGCAGAGCT
CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATG
AGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT
AAAATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATAT
TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT
CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCT
GTGAAGTGGTGGTGTGAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA
GGACACAGTGAACCTTGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAG
CAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT
TCCAGTGATGCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAATGA
TTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAATAAATA
TGTCTATCAAATACCTCTGTAGTAAATGTGAAAAAGCAAAA

FIGURE 161

MGFNLT FHL SYKFRLLLLLTLCLTVVGWATS NYFVGAIQEIPKAKEFMANFHKT LILGKGKT
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAE GKKFNRAKLLNVGYLEALKEEN
WDCFIFHDVDLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNG
FSNNYWGWWGEGDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR
VWRTDGLSSCSYKLVSV EHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

Xylose isomerase proteins:

amino acids 191-201

FIGURE 162

CGTGGGCGGGGTGCGCGCAGCGGGCTGTGGGCGCGCCCGGAGGAGCGACCGCCGCAGTTCTC
GAGCTCCAGCTGCATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGCCCCGCA**ATG**
GCCCAGGCAGTGTGGTTCGCGCCTCGGCCGCATCCTCTGGCTTGCCTGCCTCCTGCCCTGGGC
CCCGGCAGGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA
CGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG
CCCGCTGACGCCCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGCTTACTGGCAA
GATGGAGAAGGGTCTCAGCTCCACCATCCGTGTGGTTCGGCCACGTGCCCCGGGAATTCCCGG
TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG
GTCCTCCCCATCACAGAGTTCCTCGTGGGGGACCTTGTTGTCACCCAGAACACTTCCCTACC
CTGGCCAGCTCCTATCTCACTAAGACCGTCTGAAAGTCTCCTTCTCCTCCACGACCCGA
GCAACTTCTCAAGACCGCCTTGTTTCTCTACAGCTGGGACTTCGGGGACGGGACCCAGATG
GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT
CAAAGTGGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGA
CCGGGGACTTCTCCGCCTCGCTGAAGCTGCAGGAAACCCTTCGAGGCATCCAAGTGTGGGG
CCCACCCTAATTACAGACCTTCCAAAAGATGACCGTGACCTTGAACCTCCTGGGGAGCCCTCC
TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC
CTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCTGGGGACTAC
TGCTTCAGCATCCGGGCGGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT
GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA
TGTTGGCCTTCATCATGTACATGACCCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG
AACCCGGAGCCACCCTCTGGGGTCAAGTGTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT
GGAGACTCCATCTGAGTACCTGGAAATTGTTCTGTGAGAACCACGGGCTGCTCCCGCCCCTCT
ATAAGTCTGTCAAACTTACACCGTGT**AG**CACTCCCCCTCCCCACCCCATCTCAGTGTTAA
CTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCAAT
TGCCTGGGGCTGTTGGCCTGGATCATCCATCCATCTGTACAGTTCAGCCACTGCCACAAGCC
CCTCCCTCTCTGTACCCCTGACCCAGCCATTACCCATCTGTACAGTCCAGCCACTGACA
TAAGCCCCACTCGGTTACCACCCCTTGACCCCTACCTTTGAAGAGGCTTCGTGCAGGACT
TTGATGCTTGGGGTGTTCCTGTGACTCCTAGGTGGGCCTGGCTGCCCCTGCCCATTCTCT
CTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCCTCCCCCAGACAGCCCTAC
CTGTGCCAGAGAGCTAGAAAGAAGGTATAAAGGGTTAAAAATCCATAACTAAAGGTTGTAC
ACATAGATGGGCACACTCACAGAGAGAAGTGTGCATGTACACACACCACACACACACACA
CACACACACACAGAAATATAAACACATGCGTCACATGGGCATTTTCAGATGATCAGCTCTGTA
TCTGGTTAAGTCGGTTGCTGGGATGCACCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA
AGCAGCCCTGACAGGTTCTGGGCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC
GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCCTGGATGGGGGCGAGGACT
AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG
AACTTTCACTGAGGAAAAGGCCTTGACGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGG
TGGCTCACGCCTGTAATCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGATCAGGA
GATCGAGACCACCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAAATAACAAAAGTT
AGCCGGGCGTGGTGGTGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG
GTGCGAACCCGGGAGGCGGAGCTTGACAGTGAGCCCAGATGGCGCCACTGCACTCCAGCCTGA
GTGACAGAGCGAGACTCTGTCTCCA

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWPSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTNLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTFRDPGD
YCFSSIRAENIISKTHQYHKIQVWPSRIQPAVFAPFCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQMC CGPFLLET PSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

165/330

FIGURE 165

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH
FPICIFCCGCCHRSKCGMCCKT

165/330

FIGURE 166

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGC
 CTGGATCTTCCACC**ATG**TTCCCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC
 ATCTCCCTGACTGTCCTCTTACCCTCCTTCTCGTTTTTCATCATAGTGCCAGCCATTTTTGG
 AGTCTCCTTTGGTATCCGCAAACCTCTACATGAAAAGTCTGTAAAAATCTTTGCGTGGGCTA
 CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC
 GGAATCATTGCAAAGGATCCCACCTTCACTAGAAGAAGAGATCAAAGAGATTTCGTCGAAGTGG
 TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGCC
 GGAAAGGAATGGAGACCATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAAGTGA
 GAGTCCTGGAACCTGCTGAGCAGAACCAATTATAACTTCCAGTACATCAGCCTTCGGCTCAC
 GGTCTGTGGGGGTTAGGAGTGCTGATTCGGTACTGCTTTCTGCTGCCGCTCAGGATAGCAC
 TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACCTGTGGTGGGATACTTGCCAAAT
 GGGAGGTTTAAGGAATTCATGAGTAAACATGTTCACTTAATGTGTTACCGGATCTGCGTGCG
 AGCGCTGACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCT
 GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC
 ATGGTGGGTCAAGTGCACGGGGGACTCATGGGTGTGATTGAGAGAGCCATGGTGAAGGCCTG
 CCCACACGCTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA
 CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCCAGAAGGAACCTGCATC
 AATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTTGAAATTGGAGCCACAGTTTACCC
 TGGTTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA
 TGGTGACGTACCTGCTGCGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG
 CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC
 CATTGCCAGGCAGGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCTGAAGAGGGAGAAGG
 TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCAC
 AAGGACAGGAGCCGCTCC**TGA**GCCTGCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAA
 CGGGCTCAGAGCTGGAGTTGCCGCCGCCGCCCCCACTGCTGTGTCCTTCCAGACTCCAGGG
 CTCCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTCGCCGAGCCGCAGCGGGATCCCTGT
 GCACCCGGCGCAGCCTACCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA
 CGAGATGCCTTGTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAAGTCCCCA
 CCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT
 GGCGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGCGGCCACCCG
 CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTGGCTTTCGGCCTCAACATCGCCCCCAGC
 CTTGGAGCTCTGCAGACATGATAGGAAGGAAACTGTCATCTGCAGGGGCTTTTACGAAAATG
 AAGGGTTAGATTTTTTATGCTGCTGCTGATGGGGTTACTAAAGGGAGGGGAAGAGGCCAGGTG
 GGCCGCTGACTGGGCCATGGGGAGAACGTGTGTTTCGTAATCCAGGCTAACCCTGAAGTCCCC
 ATGTGATGCGCGCTTTGTTGAATGTGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGG
 GAATGGTGGTGATTCCCTACCTCACAGGGCTGTTGTGGGGATTAAAGTGCTGCGGGTGAGTGA
 AGGACACATCACGTTCAAGTACAGGCCACAAAACGGGGCACGGCAGGCCTGAG
 CTCAGAGCTGCTGCACTGGGCTTTGGATTTGTTCTTGTGAGTAAATAAACTGGCTGGTGAA
 TGA

FIGURE 167

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIVPAIFGVSTFGIRKLYMKSLLKIFAWATLRME
RGAKEKNHQLYKPVTNGIIAKDPTSLEEEIKEIRRSKSSKALDNTPEFELSDIFYFCRKGME
TIMDDEVTKRFSAAEELESWNLLSRTNYNFQYISLRLTVLWGLGVLIRYCFLLPLRIALFTG
ISLLVVGTTVVGYPNGRFEKFMSKHVHLMCYRICVRALTAIITYHDRENRPNGGICVANH
TSPIDVILASDGYAMVGQVHGGLMGVIQRAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ
DKSKLPILIFPEGTCINNTSVMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYL
LRMMTSWAIVCSVWYLPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTF
KEEQKLYSKMIVGNHKDRSRS

FIGURE 168

GCCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA
CGTCTCCTCCAGGG**ATG**GCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC
ACCTGGCAGGCCAGGCTGTTCCACCATCCTGCCCCTGGGCCTGGCTCCAGACACCTTTGA
CGATACCTATGTGGGTGTGTCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG
AAATGGCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCAGGAGACCTGGGAGGAC
AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT
CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG
GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG
GCCCTGCAGCTGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTGTTCG
AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT
TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCGGGGCTGT
GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCTG
GAAGACTCTGCTCTTGCCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCC**TGA**AAGTCCA
ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG
CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG
ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACCTCTGCTATGTGATGGGGACTTCCT
GGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTTGCAATGTGGAGACA
TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH
ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSSNTLYWELNQAVRTGGGSREL
YMRHFPPFKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS
LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

FIGURE 170

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAAT**ATG**GCTGGTTCCCCAACATGCCTCA
CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG
GTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC
TATTGTCTGGACCTTCAACACAACCCCTCTTGTCAACATACAGCCAGAAGGGGGCACTATCA
TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG
CTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT
CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG
TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG
GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC
CCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT
GCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT
GAAGGTGCTGCTGATGACCCAGATTCCCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCT
CCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAG
AGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGGAAACTCCTAACATATGCCCCCAT
TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA
TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCAC
TGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATC**TAG**ACAGCAGTG
CACTCCCCTAAGTCTCTGCTCA

FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSPKPKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW
FLKRERQEYIEKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

171/330

FIGURE 172

CTGGTTCCCCAACATGCCTCACCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC
TCTGGACCCCGTGAAAGAGCTGGTCGGTTCGGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC
CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCATAC
AGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA
GATGGAGGCTACTCCCTGAAGCTCAGCAAATGAAGAAGAATGACTCAGGGATCTACTATGT
GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG
AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG
ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT
GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG
AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCC
ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATGGTCCTCCT
GTGTCTCCTGTTGGTGCCCCCTCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC
TGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGGAA
ACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA
TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA
AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG
AATGTTATCTAGACAGCAGTGCCTCCCTAAGTCTCTGCTCAAAAAAAAAAAAAAAAAAAAA

FIGURE 173

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAA**ATGA**AAGATGCTGCTGCT
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGCACAACATATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTTGACCTATCCAAT
GCCAATCGCTGCCTCCAGGCCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAAT
TTTTCATGAAATTATTCCTCTTCCTGTTCAATAAATGATTACCCTTGCACTTAA

FIGURE 174

MKMLLLLCGLTLVCVHAEASSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ
IHVLENSLVVKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI
NEKDGETFQLMGLYGREPDLSSEDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

174/330

FIGURE 175

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAA
TGGATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTG
TCAGCTTAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTC
CCAGGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAG
AAAAAGAGCGTGCTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCA
CAGTCATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTC
ATGTGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGA
CATTCATCCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTG
GTTTCAATAAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCAC
TTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTATAGGTCTATTGCT
TGTTGGAATTCTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCTTGGCTGTC
TGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTGTGTAGTTTAAATGGGAATAAAATGTAAGTA
TCAGTAGTTTGAAAAAAAAAAAA

TCTGAGCGTTTCTGAGCCAGGGGTGACC

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA
IPATTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL
IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

177/330

FIGURE 177

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACC**ATG**AGGCT
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC
CAAGTTGCCAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAA**TAG**TGAA
AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC
CTGATCTTCACTAAAAATTGTAAAGGTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

177/330

178/330

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV
KHCTDQISFKKRLSLKKSWWK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 179

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGG
TGAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTG
GAGCAGATCCGTGGGCTGCAGACCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTC
GAACTGTGAC**ATG**GAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGG
AAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAG
CTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGG
CAAATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCGTGTACCTGAGAAGGCCATCCCAC
TCATCACTCCAGGCTCTGCCACTACTTGCT**TGA**GCACAGGACTGGCCTCCAGGGATGGCCTGA
AGCCTAACACTGGCCCCCAGCACCTCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTC
TCTCCAAGGGCAGGCTGTTAGGCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAAACTCG
CCCCACCACCCCCTCA

180/330

FIGURE 180

MERVTLAILLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK
YKSSQKQHSPVPEKAIPITPGSATTC

180/330

FIGURE 181

GTGA

FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGEKILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLLASEGS
SRLDMEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFDVAVIRERRRTLPTQGIDDFKDKAKSKTLDFIDVLLLSKDEEDG
KALSDEDIKAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

FIGURE 183

CAACAGAAGCCAAGAAGGAAGCCGTCTATCTTGTGGCGATCATGTATAAGCTGGCCTCCTGC
TGTTTGCTTTTACAGGATTCTTAAATCCTCTCTTATCTCTTCCTCTCCTTGACTCCAGGGA
AATATCCTTTCAACTCTCAGCACCTCATGAAGACGCGCGCTTAACTCCGGAGGAGCTAGAAA
GAGCTTCCCTTCTACAGATATTGCCAGAGATGCTGGGTGCAGAAAGAGGGGATATTCTCAGG
AAAGCAGACTCAAGTACCAACATTTTTTAACCCAAGAGGAAATTTGAGAAAGTTTCAGGATTT
CTCTGGACAAGATCCTAACATTTTACTGAGTCATCTTTTGGCCAGAATCTGGAAACCATACA
AGAAACGTGAGACTCCTGATTGCTTCTGGAATACTGTGTCTGAAGTGAAATAAGCATCTGT
TAGTCAGCTCAGAAACACCCATCTTAGAATATGAAAAATAACACAATGCTTGATTTGAAAAC
AGTGTGGAGAAAAACTAGGCAAACCTACACCCTGTTTATTGTTACCTGGAATAAATCCTCT
ATGTTTTGCACAAAAAAAAAAAAAAAAA

184/330

FIGURE 184

MYKLASCCLLFTGFLNPLLSLPLLDREISFQLSAPHEDARLTPEELERASLLQILPEMLGA
ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV

184/330

FIGURE 185

GAACATTTTTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT
GGGGTTGCTGGTTTTAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA
CCACCTCCGCCAGGAAGTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCCATGCCCTCCCC
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCA
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGA
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT
CAGGGGTTCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGG
GAAGAGGCCAAAGAGGCCCCAGCCGACAAGTGATCGCCCAAGCCTTACTCACCTCTCTCT
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA
CAAGCTCAGGAGGCGAATAAATGTTCAAAGTGTGTA

186/330

FIGURE 186

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG
GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

405415834660

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCTCTCGTCCAGTACCTC
GTGAACCCCGGGGTGCTCCGCACGGACCCCAGATGTCAAGAAT**ATGA**ACACGTGGCTGCTGT
TCCTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG
CTCCTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT
GAGTATGTCCCCACCCTAAGCCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC
TTACACCTCTACTTGAGTATGTCCCTAACCCCTGAGCCCCCACGCCTGGGGCCAGAGTCTTT
GTCCCCCGTGTGCGCATGTGTTCAAGGTGAGCCTCTCCCAGAAGTGAGATCATGGACAAAAA
GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCAGCAAGAAGCTGAAC
TCACGCCGAGACCTGCAGGAGTGGTGCCAGGTGCT**TGA**AGTAACAAGTTTAAATGTTTCTCAGA
GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT
CCAAAAACACAAGTAGAAATTCTAACAAATGAAATATATTACAGGCAGGTCAACCCACTAACCA
AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGT
AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT
AACAAACAACCTCCCTGCTCCTGGCACCAGCCGTTTTTGGTCATGGTGGGCCAGCTGCAAAGCG
TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC
ACGGGCAGCAGAGTGTGTCCAGGCCAGCCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA
ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTGAGGCTCTGGGCTCACCTCCATCTCCAGA
GCATCCCCTGCCTGCAGTTGTGGCAAGAACGCCAGCTCAGAATGAACACACCCACCAAGA
GCCTCCTTGTTTCATAACCACAGGTTACCCTACAAACCACTGTCCCCACACAACCCTGGGGAT
GTTTTTAAACACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA
ATTTTTTTTTTAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 188

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDLGLVHLGQLLI FHIYLSMSPTLS PRSPQGW
VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ
AQQEAELTPRPAGVVPGA

FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACCATGGCCAAAG
ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT
ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG
TGCCCCAAGCCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG
GATGGAGATACCAACACATCCACCCAGGAGGTGGTACAATACTGAGGAGTGGGGATGA
CCGTTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAACTGTGGAAG
AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA
GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA
GCGTTTGATGGAGAAGGCTTCCCTCCCCTCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA
TGTTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCT
ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAAGTGGAGCGCCTTTGCTG
CTGTTTCCTCTGTCTGTGAGGTCTCCTGGGGATGGTGGCCACATGATGTATTACAAGTC
TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGAATTATGGCTG
GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA
ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCATAGTAAGAGCTTCAAGGAAAAC
CCGAAGTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCAC
CGTGGGTCTTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG
GAGTCGACTTCTACTCCGAGCTGCGGAACAAGGGATTTCAAAGAGGGGCCAGCCAGGAGCTG
AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA
GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG
TCTCTTGAGCATGGTTTTTTAGAGGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC
CTAAGGGATTCCCTGGGTGCCACTGCTCTCTTTTCTCTACAGCTCCATCTTGTTTCACCCAC
CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTCTGGGGC
TAAACCATGGAGATAAAAAAGAAGAGTAAAAATACACTTCCCCGACCTTAAGGATCTGAAA

190/330

FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFGVTQKVPKPLCEKGLAAKCFDMP
VSLDGDNTSTSTQEVVQYNWETGDDRFSSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR
GEKGLLEFATLQGPGCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL
PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSPDCQLGSRRLTTTCLE
LWLGLLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

190/330

FIGURE 191

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAG**ATGT**GGTTACCCCTTGGTCTCCTG
TCTTTATGTCTTTCTCCTCTTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA
GCAGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG
TTATGGATGTTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT
GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG
AGATGGCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCCTGCCCTATTCCTCCTCCCAA
GTCTGTTCTCTTATTGTCAACCTCAGCACAAACAGGCTGGCGCCAATGGCATTACAGAGAAAG
CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT
AGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTAG
GACATTCGCCCCCTGTGTGCCACCAAACCAGGACTTTCCCCTTGGCTTGGCATCCCTGGCTCT
CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT
ATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA
GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCCTGAGGAAAAACCAA
GGGAAGCAACAGGAACCTTCTGCAACTGGTTTTTATCGGAAAGATCATCCTGCCTGCAGATGC
TGTTGAAGGGGCACAAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA
ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAACTCTTTATTACTTTGGG
AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG
AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC
CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA

GenBank J01491

192/330

FIGURE 192

MWLPLGLLSLCLSPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQGSME
HRNHLCFCDLYDRATSPPLKCSLL

192/330

FIGURE 193

GTAGCGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCTCGGGTCGTGGAGCCAGGAGCGACGTCA
 CCGCCATGGCAGGCATCAAAGCTTTGATTAGTTTTGTCTTTGGAGGAGCAATCGGACTGATGTTTTGATGCTT
 GGATGTGCCCTTCCAATATACAACAAATACTGGCCCTCTTGTCTATTTTTTACATCCTTTCACCTATTCC
 ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACTTGCCATCTTTC
 TTACAACGGGCATTGTCTGTCTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGA
 GCTTGTGCACTTGTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAG
 CAATGACGACTTCAGCTGGCAGCAGTGGTGAAAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCATTT
 GTTGGCCATTACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA
 GGTGCTCCCTTCTCACTTTTATTGTAAGCATACTATTTTACAGAGACTTGCTGAAGGATTTAAAGGATTTTCT
 CTTTTGGAAAAGCTTGACTGATTTTACACTTATCTATAGTATGCTTTTTGTGGTGTCTGCTGAATTTAAATAT
 TTATGTGTTTTTCTGTTAGGTTGATTTTTTTTTGGAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCA
 TTTGCATTGGTTAGGAATTCAGAATTCGCCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAAAATTATT
 TAGCCTCCATTATTACAAAAAATTATAAAAAATAAGTTTTTTCAGTCAGTCAGGATGACATCACTCCCAATGTTATG
 CAGACATACAGACGGTTGGCATACGTTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG
 GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTGTGGTTTTACTGGTAGACAGATGTTTTGTGGATTG
 AAAATTATTTTATGGAATTGCTACAGAGGAGTGCTTTTCTTCAATTGTTAGAAGAATTTATGTTAAACTTTA
 AGGTAAGGGTGTA AAAACATTTTTGAGATAAGGTTTTTATTTATGTTTATTATTGTTAGAGTGAGTTGCAATGT
 GGGAAAGAAATGACATTGAAATTCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC
 AACCTTTTCATGTTTTACCCTGTTAAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC
 ATCATATATGCCAGAAAACCTTCTCTGCTTCTCTCTTTTGACTTATTTGGTATGTTGTATATATTACATAAAA
 TAACTTTTCAAATATAGTTTAATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTACATT
 CAGAGTGCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAAGTTGTTAGTCTTACAGATAATTGATGCA
 TTAACAGTTTTAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAAG
 TATTTTTAAGACAAGTTTCTGTATACCTCTGAAGTGTGTTGATTTTGAGTTCATCATGATAGATCTGCTGTTT
 CCTTATAAAAGGCATTTGTTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT
 ACCTGACCAAAAAATTCCCAGTAACCAAGGCATGATCAATTTATAGTGGTCTGTTTACATCTAATAATTATCAGGA
 CTTTTTTCAGGAGTGGGTTATAAAAACATTCAAGTTGGTCTGACAGTATTTGTTAAGGATATTTGTTTGTATG
 TTTATTCAGTATACCTACATAAAAAATTATTCGCCATCAGCCAAAACCTCAGTAATCATGACAGCTGTCTGTTGT
 TTTATGAAGTTTTATTTCTCAAGAAAATGGGAATAAATTTGGGATTTGTTTCTGCTTTTTACTAAAGATGCCTAA
 AGCCACAGGTTTTATTGCCTAACTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG
 GCGTGTGGCTGGAGCCTTCCCACTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTTCAAGAGGAA
 GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACCTTTGTGCTTGTGATCTACTGGACTTT
 TTTTTTGCAGGAAGTGCATTCTCTGGTCTTCCCTATTTTCTGTTCTGGATGTCAGTGCAGTGCAGTGCCTACTG
 TTTTATCCACTTGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT
 GTGTCTTTGACCTTGATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT
 TTTCCATAGAATATGCACTGATACAACATTACCAATCTTCTATGGAAAGAAAACTTTTGATGATGAAACAATAA
 AGATTTTTAAATATCTATTTTAAAAA

[illegible]

FIGURE 196

MDFLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLLSLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVYAFHLCEDGAEPTSG
HLLSAVTNRSDLGPPASSATTADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH
IEGALVINEYGSCTCHQQPARECEV

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE
GHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA
PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVVRVRSISALGRTFTRDEDLAVFLASR
AGRLRFHGPGLSVGPEDCADPSGCVCGNAEAQPWICAALLQP

FIGURE 199

ATCGCATCAATTGGGAGTACCATCTTCCTC**ATG**GGACCAGTGAAACAGCTGAAGCGAATGTT
TGAGCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTT
CTGCCTTTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCA
TTGACGTGGTACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTTTTGC
CGTGTGTCTTGCA**TAA**TTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAG
CTGGTGGACAGTTTTGTAACCTCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCT
TGCAGCAATGTGTTGCTTGTGATTTCGAACATTTGAGGGTTACTTTTGAAGCAACAATACAT
TCTCGAACCTGAATGTCAGTAGCACAGGATGAGAAGTGGGTCTGTATCTTGTGGAGTGGAA
TCTTCCTCATGTACCTGTTTCCTCTCTGGATGTTGTCCCACTGAATCCCATGAATACAAAC
CTATTCAGCAACAGCAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 201

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGCGTTCCCTCCAGTCACCCTCCCGCCGTTACCCGCGGGCGCGC
 CCGAGGGAGTCTCCTCCAGACCCTCCCTCCCGTTGCTCCAACTAATACGGACTGAACGGATCGCTGCGAGGGT
 GGGAGAGAAAAATTAGGGGGAGAAAGGACAGAGAGAGCAACTACCATCCATAGCCAGATAGATTATCTTACACTG
 AACTGATCAAGTACTTTGAAAATGACTTCGAAATTTATCTTGGTGTCTTCATACTTGCTGCACTGAGTCTTTC
 AACCACCTTTTCTCTCCAAGTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT
 TATATAAGTTCCAACGCCCCATTTTCATTATATTATGAAATATGGTGTTACAGTGAAGCAAGTTACTAATGTT
 TTTATTACAAAAACCTACCCTAACCATTTACTTTGGTAACTGGCCTCTTTGCAGAGAATCATGGGATTGTTGC
 AAATGATATGTTTGATCCTATTCGGAACAAATCTTCTCCTTGATCAGATGAATATTTATGATTCCAAGTTTT
 GGAAGAAGCGACACCAATATGGATCACAACCCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGA
 ACAGATGTAAAAATACATAAGCGCTTTCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAG
 AGTTGCCAAAAATTGTTGAATGGTTTACGTCAAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTG
 ATGACATGGGCCACCATTGGGACCTGACAGTCCGCTCATGGGGCCTGTCATTTTCAAGATATTGACAAGAAGTTA
 GGATATCTCATACAAATGCTGAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATCATCACAAGTGATCATGG
 AATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTACCTGGATAAAGACCACTATACCCTGATTG
 ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAAATTTGATGAAGTCTATGAAGCACTAACTCAGCT
 CATCCTAATCTTACTGTTTACAAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTC
 ACCAATCATAGCAGTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACC
 ACGGTTACGATAATGCGTTAGCAGATATGCATCCAATATTTTACGCCATGGTCTGCCTTCAGAAAGAATTTCT
 TCAAAAGAAGCCATGAACCTCCACAGATTTGTACCCACTACTATGCCACCTCCTCAATATCACTGCCATGCCACA
 CAATGGATCATTTCTGGAATGTCCAGGATCTGCTCAATTCAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA
 CTATACTCCTCCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTCTAACCTTATTTCATAGGGGTC
 TCTCTTGGCAGCATTATAGTGATTGTATTTTGTATTTTTCATTAAAGCATTAAATTCACAGTCAAATACCTGC
 CTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCCTAATGTTACTTTGAAGTGGATTTCATA
 TTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAATTCGGGAAACCAGTTCCAAACATCTGC
 AGAAACCATTAAGCAGTTACATATTTAGGTATACACACACACACACACACATACACACACACGGACCAAA
 ATACTTACACCTGCAAAGGAATAAAGATGTGAGAGTATGTCTCCATTGTTCACTGTAGCATAGGGATAGATAAG
 ATCCTGCTTTTATTTGGACTTGGCGCAGATAATGTATATATTTAGCAACTTTGCACTATGTAAAGTACCTTATAT
 ATTGCACTTTAAATTTCTCCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAAC
 TTGATTGAAATGACAACCTTTTGCACCCATGTCACAGAATACTTGTTACGCATTGTTCAAACCTGAAGGAAATT
 TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTGA
 AAATTAAATGTGATAACCTTTGAACCTTTGAATTTTGGAGATGTATTCCCAACAGCAGAATGCAACTGTGGGCAT
 TTCTTGTCTTATTTCTTTCCAGAGAACGTGGTTTTCATTTATTTTCCCTCAAAAGAGAGTCAAATACTGACAG
 ATTTCGTTCTAAATATATTGTTTCTGTCTATAAAATTATTGTGATTTCCCTGATGAGTCATATTACTGTGATTTTCA
 TAATAATGAAGACACCATGAATATACTTTTCTTCTATATAGTTTCAGCAATGGCCTGAATAGAAGCAACCAGGCA
 CCATCTCAGCAATGTTTCTCTTGTGTAATTATTTGCTCCTTTGAAAATTAAATCACTATTAATTACATTAA
 AAATCAAATTGGATAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW
ITNQ RAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTELIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

FIGURE 203

GGATTTTTGTGATCCGCGATTCTGCTCCACGGGCGGGACCTTTGTAAGTGCGGGAGGCCCAG
 GACAGGCCCCACCTGCGGGGCGGGAGGCAGCCGGGGTGAGGGAGGTGAAGAAACCAAGACGC
 AGAGAGGCCAAGCCCCCTTGCCTTGGGTACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA
 CCAGATCCAGAGGCAACAGGGACATGGCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC
 AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA
 CTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAATGAAGAGGAGGAGGAGGAGGAGG
 AGCAGCCACCACCCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCCTGACGTTGCC
 CCTGCCCCCTGGCCCCGCACCCAGGGCCCCCCTTGAAGTTCAGGGGCATGTTGAGGAACTGTT
 CAGCTCCCACAGGTTTCAGGTCATCATCATCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGC
 TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG
 GTATTCCACTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCTTTAAATT
 ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGG
 TGGTCTCATTATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC
 CTGCTGATTCTGCTCCGGCTGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT
 TAAGACACGTTGAGAACGGCAACTCTTAAGGTTAAACAGATGAATGTACAATTGGCCGCCA
 AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCTGGACTTGATGAGTTTGCTGTATC
 AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT
 CACACAGCCACCGTGAAAGTCCTGGAGTAAAATGTGCTGTGTACAGAAGAGAGAGAAGGAAG
 CAGGCTGGCATGTTCACTGGGCTGGTGTACGACAGAGAACCTGACAGTCACTGGCCAGTTA
 TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC
 AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEEQPPPTPV
SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSSHRFQVIIICLVVLDALLVLAELILDL
KIIQPDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFRLLSSFTTSLRSWMPVVVVVSFILDI
VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLRLKQMNVQLAAKIQHLEFS
CSEKPLD

FIGURE 205

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCAGTGGAGCACCCAGCAGGCCGCCAAC**AT**
GCTCTGTCTGTGCCTGTACGTGCCGGTCAATCGGGGAAGCCCAGACCGAGTTCCAGTACTTTG
 AGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCCTCC
 CAGGAATTTCTCCACCTACCGCCAGTGGAAGCAGAAAATTGTACAAGCTGGAGATAAGGACCT
 TGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGA
 GGCTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATG
 CAGTCCCTGCGGGACTTGGGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAG
 CATGGATAAAAAACGGCACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCC
 ACCCCGTGGAAAACATCCCCGAGATCATCTCTACTGGAAGCATTCCACGATCTTTGATGTG
 GGTGAGAATCTAACGGTCCCGGATGAGTTACAGTGGAGGAGAGGCAGACGGGGATGTGGTG
 GAGACACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGG
 ACAGGCTCAAGGTGCTCATGCAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGT
 GGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAA
 CGTCTCAAATTTGCCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAGATCAAGCGCC
 TTGTTGGTAGTGACAGGAGACTCTGAGGATTCACGAGAGGCTTGTGGCAGGGTCTTTGGCA
 GGGGCCATCGCCCAGAGCAGCATCTACCCAATGGAGGTCTGAAGACCCGGATGGCGCTGCG
 GAAGACAGGCCAGTACTCAGGAATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGG
 TGGCCGCTTCTACAAAGGCTATGTCCCCAACATGCTGGGCATCATCCCCTATGCCGGCATC
 GACCTTGCAGTCTACGAGACGCTCAAGAATGCCTGGCTGAGCAGCATATGCAGTGAACAGCGC
 GGACCCCGCGTGTGTTGTGCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGG
 CCAGCTACCCCTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCT
 CCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCCTGCGGACCGAGGGGGCCTTCGGGCT
 GTACAGGGGGCTGGCCCCCAACTTCATGAAGGTCACTCCAGCTGTGAGCATCAGCTACGTGG
 TCTACGAGAACCTGAAGATCACCTGGGCGTGAGTCGCGG**TG**ACGGGGGGAGGGCCGCCCG
 GCAGTGGACTCGCTGATCCTGGGCCGACGCTGGGGTGTGCAGCCATCTCATTCTGTGAATG
 TGCCAACACTAAGCTGTCTCGAGCCAAGCTGTGAAAACCCTAGACGCACCCGCAGGGAGGGT
 GGGGAGAGCTGGCAGGCCCAGGGCTTGTCTGTGACCCAGCAGACCCTCCTGTTGGTTCC
 AGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCAGCAGGCTCCGGGCTCACATGTGTAA
 GGACAGGACATTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCTTAGT
 TCTTCCATTTACCCCTTGACGCCAGCTGTTGGCCACGGGCCCTGCCCTCTGGTCTGCCGTGC
 ATCTCCCTGTGCCCTCTTGTGCTGCTGCTGCTGAGGTAAGGTGGGAGGAGGGCTACAG
 CCCACATCCCACCCCTCGTCCAATCCCATAATCCATGATGAAAGGTGAGGTCACGTGGCCT
 CCCAGGCCTGACTTCCCAACCTACAGCATTTGACGCCAACTTGGCTGTGAAGGAAGAGGAAAG
 GATCTGGCCTTGTGCTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATGCT
 TGGGAGTGACGGGGGCTCGGGCTGCTGGCCTGGCTGCAACAGAAGGCAAGTCTGGGGCTCA
 TGGTGCTCTGAGCTGGCCTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAAACCTCACTG
 TCCCCACTGTGGCATGAGGGCAGTGGAGCACCATGTTTGAAGGCGAAGGGCAGAGCGTTTGT
 GTGTTCTGGGGAGGGAAGGAAAAGGTGTTGGAGGCCTTAATTATGGACTGTTGGGAAAAGGG
 TTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
 GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTTCTGTCCAACC
 CCAGCAGGGGCGCAGCGGGACCAGCCCCACATTCACCTTGTGTCACTGCTTGGAACCTATTT
 ATTTTGTATTTATTTGAACAGAGTTATGTCTTAATTTATATAGATTTGTTTAATTAATA
 GCTTGTCAATTTCAAGTTCAATTTTTATTTCATATTTATGTTTCATGGTTGATTGTACCTTCCC
 AAGCCCGCCAGTGGGATGGGAGGAGGAGAAGGGGGGCCCTTGGGCCGCTGCAGTCACAT
 CTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAGGCAGCAGCCCTG
 GCTCCTTTTCTTTTGGCAGGTTGGGGAAGGGCTTGCCCCCAGCCTTAGGATTTTCAAGGTTTGA
 CTGGGGGGCGTGGAGAGAGAGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTT
 CTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGAAGTGCCTCT
 CACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGTCCTCC
 CAGCCTTCTGCTGCCCTTGCTTAACAATGCCGGCCAACCTGGCGACCTCACGGTTGCACTTCC
 ATTCCACCAGAATGACCTGATGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAAATT
 GTTATATATGAACATATAACTGGAGTCGTCAAAAAGCAAATTAAGAAAGAATTGGACGTTAG
 AAGTTGTCAATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAAAAAAAAAAAAAA
 AA

FIGURE 206

MLCLCLYVPVIGEAEQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD
LDGQLDFEEFVHYLQDHEKKRLRVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILK
SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW
WRHLVAGGGAGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI
NVLKIAPESAIFMAYEQIKRLVGSQETLRIHERLVAGSLAGAIQSSIYPMEVLKTRMAL
RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIIPYAGIDLAVYETLKNWLQHYAVNS
ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFG
LYRGLAPNFMKVI PAVSISYVYENLKITLGVQSR

Important features:**Signal peptide:**

amino acids 1-16

Transmembrane domain:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation site.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 208

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI
KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKKNVQLTD
AGTYKCYIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTV VWASQVD
QGANFSEVSNTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTESEIKRR
SHLQLLNSKASLCVSSFFAISWALLPLSPYMLK

FIGURE 209

GAATTTGTAGAAGACAGCGGCGTTGCC**ATG**GCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGT
GCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGC
AGCGGGTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTG
AATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCAC
CTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACTCTCTGGCTATCATCTTCA
CACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGC
GAGTGCGGGACGCAGCTCTGTGGATCTCGACATACCTGTGTTAGTTCCTTCCCAGAACCCAT
CTCCCCAGAGTGGGTGAGGACACGGCCTTTTCCCATCCTGCCCTTTCCTCTGCAGCTGTTTT
GCTTCCTTGTGGCCATCAGAGTTCCCTTCCCCTGGACAGTCTGGAGAAAGACAGAGGCTGGG
GTTTGGGAT**TGA**AGACCAGACCCCATCTGAGCCCTTCCTCCAGCCCTGTACCAGCTCCTACT
GGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCAT
GGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGC
AACAGTCTACCCTTTGAGTGGGCGGAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCCT
GGGCCATGAAGTGCTGGCAGTGAGCGGATGGACCTAGCACTTCCCCTCTCTGGCCTTAGCTT
CCTCCTCTCTTATGGGGATAACAGCTACCTCATGGATCACAATAAGAGAACAAGAGTGAAAG
AGTTTTGTAACCTTCAAGTGCTGTTTCAGCTGCGGGGATTTAGCACAGGAGACTCTACGCTCA
CCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCCCTGCTAACATCTCAGGCTCCCAGCCCA
GCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACT
CCAGCTGCATGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTG
GGACTCAGTCAGAGCGCTTTGGCTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGA
GTGGCAGATGCAGGAAATGAGCTGTCTATTAGCCTTGCTGCCCCACCCATGAGGTAGGCAG
AAATCCTCACTGCCAGCCCCCTTTAAACAGGTAGAGAGCTGTGAGCCCCAGCCCCACCTGAC
TCCAGCACACCTGGCGAGTAGTAGCTGTCAATAAATCTATGTAAACAGACAAAAAAAAAAAAA
AAA

FIGURE 213

GGCCTCGGTTCAAACGACCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCA
GGGCTTGCCTCACTGGCCACCCTCCCAACCCCAAGAGCCCAGCCCC**ATG**GTCCCCGCCGCCG
GCGCGCTGCTGTGGGTCTGCTGCTGAATCTGGGTCCCCGGGCGGCGGGGGCCCAAGGCCTG
ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG
CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA
ATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC
ACGGTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAAACGAGGAGGATGGGTCTTCAGAAGA
GGGGGTGTGATTAATGCCGGAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCA
ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG
ACTTCAAGCCTGCCGCGCTCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC
CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCAACCTCACCCACAG
CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC
TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCCGGGCGCCT
TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGACCTATCAACAATGTC
CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT
GCCTCTCAGAGCACCACCAGTACCAGGACCACCACTACCCCTTCCCCACCATCCACCTCAG
AAGCAGTCCCAGCCTGCCACCCGCCAGCCCCTGCCAGCCCTGGCTTTTTTGGAACGGGTCA
GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTACAGAGATGCAACCAATA
GACAGAAACCAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT
CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAAACAAACTGGAAAA
CACAA

FIGURE 214

MVPAAGALLWVLLLNLGPRAAGAQGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI
ILEDENDAMADADRLAGPAAAELLAATVSTGFSRSSAINEEDGSSEEGVVINAGKDSTSREL
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP
SPSPTAMPSPEDLRLVLMWPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC
TYQQCPCNRLREECPLDTSLCTDTNCASQSTTSTRTTTTTFPTIHLRSSPSLPPASPCPALA
FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

CCGCGGTCGACCCACGCGTCCGGGGAGAAAGGATCGCCGGCCTGGCGGCGCGGTTGGTCCTGCTAGCTGGGGCA
GCGGCGCTGGCGAGCGGCTCCCAGGGCGACCGTGAGCCGGTGTACCGCGACTGCGTACTGCAGTGCGAAGAGCA
GAATGCTCTGGGGGCGCTCTGAATCACTTCCGCTCCCGCCAGCCAATCTACATGAGTCTAGCAGGCTGGACCT
GTCGGGACGACTGTAAGTATGAGTGTATGTGGGTCAACCGTTGGGCTCTACCTCCAGGAAGGTACAAAAGTGCCT
CAGTTCCATGGCAAGTGGCCCTTCTCCGGTTCTGTCTTTCAAGAGCCGGCATCGGCCGTGGCCCTCGTTTCT
CAATGGCCTGGCCAGCCTGGTGATGCTCTGCCGTAACCGACCTTCGTGCCAGCCTCTCCCCATGTACCA
CCTGTGTGGCCTTCGCCTGGGTGTCCCTCAATGCATGGTTCTGGTCCACAGTCTTCCACACAGGAGACTGAC
CTCACAGAGAAAATGGACTACTTCTGTGCCTCCACTGTCATCCTACACTCAATCTACCTGTGCTGCGTCAGGAC
CGTGGGGCTGCAGCACCCAGCTGTGGTCAGTGCCTTCCGGGCTCTCCTGCTGCTCATGCTGACCGTGCACGTCT
CTACCTGAGCCTCATCCGCTTCGACTATGGCTACAACCTGGTGGCCAACGTGGCTATTGGCCTGGTCAACGTG
GTGTGGTGGCTGGCCTGTGCCTGTGGAACACGGCGGGCTGCCCTACGTCGCGCAAGTGCCTGGTGGTGGTCTT
GCTGCTGCAGGGGCTGTCCCTGCTCGAGCTGCTTGACTTCCCACCGCTCTTCTGGGTCTGGATGCCATGCCA
TCTGGCACATCAGCACCATCCCTGTCCACGTCTCTTTTTCAGCTTTCTGGAAGATGACAGCCTGTACCTGCTG
AAGGAATCAGAGGACAAGTTCAAGCTGGACTGAAGACCTTGGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCC
GCCCTGCTGGCCTCCCTTCTCCCTCAACCCTTGAGATGATTTTCTCTTTTCAACTTCTTGAAGTGGACATGA
AGGATGTGGGCCAGAATCATGTGGCCAGCCACCCCTGTTGGCCCTCACCAGCCTTGGAGTCTGTTCTAGGG
AAGGCTCCAGCATCTGGGACTCGAGAGTGGGCAGCCCTCTACCTCCTGGAGCTGAACTGGGGTGGAACTGA
GTGTGTTCTTAGCTCTACCGGGAGGACAGCTGCCTGTTCTCTCCACAGCCTCTCTCCCATACCCAGCTG
CCTGGCTGGGTCTGAAGCCCTCTGTCTACCTGGGAGACCAGGAGACCAGGCTTAGGATACAGGGGCTCC
CTTCTGTTACCACCCCCACCCTCCTCCAGGACACCACTAGGTGGTGTGCTGGATGCTTGTCTTTGGCCAGCCAA
GGTTCACGGCGATTCTCCCATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTACCCTGACC
GTTGCCCTAGCCAGGTTCCCAGGAGGCTCACCATACTCCCTTTTCAAGGCCAGGGCTCAGCAAGCCAGGGCA
AGGATCCTGTGCTGCTGTGTTGAGAGCCTGCCACCGTGTGTGCGGAGTGTGGGCCAGGCTGAGTGCATAGG
TGACAGGGCCGTGAGCATGGGCTGGGTGTGTGTGAGCTCAGGCTAGGTGCGCAGTGTGGAGACGGGTGTGT
CGGGGAAGAGGTGTGGCTTCAAAGTGTGTGTGTGCAGGGGGTGGGTGTGTAGCGTGGGTAGGGGAACGTGT
TGCGCGTGTCTGGTGGGCATGTGAGATGAGTGACTGCCGGTGAATGTGTCCACAGTTGAGAGGTTGGAGCAGGAT
GAGGGAATCCTGTCAACATCAATAATCACTTGTGGAGCGCCAGCTCTGCCCAAGACGCCACCTGGGCGGACAGC
CAGGAGCTCTCCATGGCCAGGCTGCCTGTGTGCATGTTCCCTGTCTGGTGCCCCCTTGGCCGCCCTCTGCAAC
CTCACAGGGTCCCCACACAACAGTGGCCCTCAGAAGCAGCGCCCTCGGAGGCAGGGAAGGAAAATGGGGATGGC
TGCGGCTCTCTCCATCCTCCTTTTCTCCTTGGCTTCGATGCGCTGGCCTTCCCCCTCAAACCTCCATTCCCCT
GCTGCCAGCCCCCTTGGCATAGCCTGATTTTGGGGAGGAGGAAGGGGCGATTGAGGGAGAAGGGAGAAAGCT
TATGGCTGGGTCTGGTTTCTTCCCTTCCCAGAGGGTCTTACTGTTCCAGGGTGGCCCCAGGGCAGGCAGGGGCC
ACACTATGCCTGTGCCCTGGTAAAGGTGACCCCTGCCATTTACCAGCAGCCCTGGCATGTTCTGCCCCACAGG
AATAGAATGGAGGGAGCTCCAGAACTTTCCATCCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTG
CTCTGCCCTGACCCCTTGTCCCTCTTTGAGGGAGGGGAGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTG
GCTGCGCTAGCTTCTTTGTATACGAAAACTTTTAAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAA
TTCCAAGCCTCAAAAAAAAAAAAAAAAAA

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPESSTKETERKETKAEHEEL
 DAEVLEV FHP THEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN
 TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
 RLINKFNSSSSSLEEKIAALFDLEYVYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF
 VLGAAFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
 GGLQVLR TLVQEKGTEVLAVRVVTLTYDLVTEKMFEEEEAELTQEMSPEKLQQYRQVHLLPG
 LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGR TLASLQAEYQVLAS
 LELQDGEDEGYFQELLGSVNSLLKELR

Important features:**Signal peptide:**

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

[illegible]

FIGURE 220

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIIILVAGAFFWLVSLLLASVWVFILVHVTDR
SDARLQYGLLI FGAAVSVLLQEVRFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS
FGIISGVFSVINILADALGPGVVGIHGDSPIYYFLTSAFLTAIIILLHTFWGVVFFDACERRR
YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCKD

FIGURE 221

AAGCTGGTTTAAAGGAAGCAGAGGAGGGTTAGATTCGTTGAGTGAGGACGGAAGATCAACCCA
TTTCCATTCCGCCAGATGGCCTATGTTTCTGGTCTCTCCCTTCGGNATCATCAGTGGTGTNT
TNTCTGTTATCAATATTTTGGCTGATGCANTTGGGCCAGGTGTGGTTGGGATCCATGGAGAC
TCACCCTATTANTTCCTGANTTCAGCCTTTNTGACAGCAGCCATTATCCTGCTC

2025-11-23 14:00:00

FIGURE 222

GACCGACCGTTCAGATGCCCGGTTCCAGTACGGCTTCCTGATTTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTNTCTGTTATCAATATTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCCCTATTAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA
GAGGAG

FIGURE 223

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCGGTTTCCCCCTTTCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTCCTGTCTGGTCAGGCCCCCACCCCCCTTCCCACNTG
ACCAGCCATGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTGCGGTTTCGGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGAC
CGACCGGTCAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTTGGTGCTGCTGTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTTCCCNNTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNCTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCCC
CAGNTGGCGCGCCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCCTTCCCACCTGA
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTCGGGCCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGA
CCGACCGGTCAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

FIGURE 225

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCCAGGAGGAGGCAG
 TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCC
 TACCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGG
 TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCT
 AGACTCCTATCTTCTGAATTCTATAGTGCCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCCGT
 CCTTGTGGTTCCTCTCTACCTGGGGAAATAAGGTGCAGCGGCCATGGCTACAGCAAGACCCC
 CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTT
 CTCGCCAACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA
 GGACCTGGGAGCTGGGGCCGGGGAAGACGCCCGGTGGATGACAGCAGCAGCCGCATCATCA
 ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC
 CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG
 GAAGAAAGTTTTTCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC
 AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCTGGCTACTCCCACCCTGGCCACTCT
 AACGACCTCATGCTCATCAAACCTGAACAGAAGAATTCGTCCCCTAAAGATGTCAGACCCAT
 CAACGTCTCCTCTCATTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAA
 CCAAGAGCCCCCAAGTGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGT
 CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA
 CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC
 TGCAGGGACTCGTGCTCCTGGGGAGATTACCCTTGTCGCCGGCCCAACAGACCGGGTGTCTAC
 ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTTGAGTCAT
 CCCAGGACTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAG
 ACCCTCATTCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCCTGACCCCATGTCT
 CCTGGACTCAGGGTCTGCTTCCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG
 GAACAATTTCCAAAACCTGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCAT
 CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA
 CTGAGAAGTGGAATAAAAAA

FIGURE 226

MATARPPWMWVLCALITALLLGVTEHVLANNNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD
SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS
PVYESGQQMFQGVKSIHPHGYSHPGHSNDLMLIKLNRRI RPTKDVRPINVSSHCP SAGTKCL
VSGWGTTKSPQVHF PKVLQCLNISVLSQKRCE DAYPRQIDDTMFCAGDKAGR DSCQGD SGGP
VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

ATGGTCAACGACCGGTGGAAGACCATGGCGGCGCTGCCCAACTTGAGGACCGGCCGCGCGA
CAAGCCGCAGCGGCCGAGCTGCGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC
TGCTGGCTGTAGCTGTACCGGTGCCGTGCTCTTCCTGAACCACGCCACCGCGCCGGGCACG
CGCCCCCACCCTGTCTGCTAGCACTGGGGCTGCCAGCGCCAACAGCGCCCTGGTCACTGTGGA
AAGGGCGGACAGCTCGCACCTCAGCATCTCTCATTGACCCGCGCTGCCCGACCTCACCAGACA
GCTTCGCACGCCTGGAGAGCGCCCAGGCCTCGGTGCTGCAGGCGCTGACAGAGCACCAAGGCC
CAGCCACGGCTGGTGGGCGACCAGGAGCAGGAGCTGCTGGACACGCTGGCCGACCAGCTGCC
CCGGCTGCTGGCCCCGAGCCTCAGAGCTGCAGACGGAGTGATGGGGCTGCGGAAGGGGCATG
GCACGCTGGGCCAGGGCCTCAGCGCCCTGCAGAGTGAGCAGGGCCGCTCATCCAGCTTCTC
TCTGAGAGCCAGGGCCACATGGCTCACCTGGTGAACCTCCGTGAGCAGCATCCTGGATGCCCT
GCAGAGGGACCGGGGGCTGGGCCGGCCCCGCAACAAGGCCGACCTTCAGAGAGCGCCTGCC
GGGGAACCCGGCCCCCGGGGCTGTGCCACTGGCTCCCGGCCCCGAGACTGTCTGGACGTCCTC
CTAAGCGGACAGCAGGACGATGGCGCTCTACTCTGTCTTTCCACCCACTACCCGGCCGGCTT
CCAGGTGTACTGTGACATGCGCACGGACGGCGCGCGCTGGACGGTGTTTCAGCGCCGGGAGG
ACGGCTCCGTGAACCTTCTCCGGGGCTGGGACGCGTACCGAGACGGCTTTGGCAGGCTCAC
GGGGAGCACTGGCTAGGGCTCAAGAGGATCCACGCCCTGACCACACAGGCTGCCTACGAGCT
GCACGTGGACCTGGAGGACTTTGAGAATGGCACGGCCTATGCCCGCTACGGGAGCTTCGGCG
TGGGCTTGTTCTCCGTGGACCCTGAGGAAGACGGGTACCCGCTCACCGTGGCTGACTATTCC
GGGCACTGAGGCGACTCCCTCCTGAAGCACAGCGGCATGAGGTTCCACCAAGGACCGTGA
CAGCGACCATTAGAGAACAACATGTGCCGCTTCTACCGCGTGCCCTGGTGGTACCGCAACT
GCCACACGTCCAACCTCAATGGGCAGTACCTGCGCGGTGCGCACGCCTCCTATGCCGACGGC
GTGGAGTGGTCTCTGGACCGGCTGGCAGTACTCACTCAAGTTCTCTGAGATGAAGATCCG
GCCGGTCCGGGAGGACCGCT**TAGACT**TGGTGCACCTTGTCCTTGGCCCTGCTGGTCCCTGTCCG
CCCATCCCCGACCCCACTCTCTCTTCTGTAATGTTCTCCACCCACTGTGCCTGGCGGAC
CCACTCTCCAGTAGGGAGGGGCGGGCCATCCCTGACACGAAGCTCCCTGGGCCGGTGAAGT
CACACATCGCCTTCTCGCCGTCCCCACCCCTCCATTTGGCAGCTCACTGATCTCTTGCCTC
TGCTGATGGGGGCTGGCAAACCTTGACGACCCCAACTCCTGCCTGCCCCCACTGTGACTCCGG
TGCTGTTTTGCCGTCCCCTGCCAGGATGGTGGAGTCTGCCCCAGGCACCCCTGCCCCCTCC
GGCCAAATACCCCGCATTATGGGGACAGAGAGCAGGGGCGAGACAGACCCCTGGAGTCCCTC
CTAGCAGATCGTGGGGAATGTGAGGTCTCTCTGAGGTGAGGTCTGAGGCCAGTATCCTCCAG
CCCTCCCAATGCCAACCCCAACCCCGTTTCCCTGGTGCCAGAGAACCCACCTCTCCCCAA
GGGCTCAGCCTGGCTGTGGGCTGGGTGGCCCCATCCTACCAGGCCCTGAGGTCAGGATGGG
GAGCTGCTGCCCTTTGGGGACCCACGCTCCAAGGCTGAGACCAGTTCCTGGAGGCCACCCAC
CCTGTGCCCGGCGAGGCTGGGTCTGCAGTCTCTTACCTGCTGTGCCACCTGCTCTCTG
TCTCAAATGAGGCCCAACCCATCCCCACCCAGCTCCCGCCGCTCCTCCTACCTGGGGCAGC
CGGGGCTGCCATCCCATTTCTCCTGCCTCTGGAAGGTGGGTGGGGCCCTGCACCGTGGGGCT
GGACTGCGCTAATGGGAAGCTCTTGTTTTCTGGGCTGGGGCTAGGCAGGGCTGGGATGAG
GCTTTGACAAACCCCAACCAATTTCCAGGGACTCCAGGGTCTTGAGGCTCCTCCAGGAGG
GCCTTGGGGGTGATGACCCCTTCCCTGAGGTGGCTGTCTCATGAGGAGGCCAACCCCTTGCC
ATTGACCGTGGCCACCTGGACCCAGGCCAGGCCCGGCCCGGCGAGTGGTCAAGGGACAGGGA
CCACCTCACCGGGCAAATGGGGTCGGGGGGGACTGGGGCACCAGACCAGGCACCACCTGGACA
CTTTCTTGTTGAATCCTCCCAACACCCAGCAGCTGTATCCCACTCCTTGTTGTGCACACA
TGCAGAGGTGAGACCCGAGGCTCCAGGACGAGCAGGCACACAGGGCAGGGCTGGAGCCGGG
TCCTCAGCTGTCTGCTCAGCAGCCCTGGACCCGCGTGCCTTACGTGAGGCCAGATGACAGG
CGGCTTTTCCAAGGCCTCCTGATGGGGGCTCCGAAAGGGCTGGAGTCAGCCTTGGGGAGCT
GCCTAGCAGCCTCTCCTCGGGCAGGAGGGGAGGTGGCTTCTCCAAAGGACACCCGATGGCA
GGTGCTAGGGGCTGTGGGGTTCCGTTCTCCCTTCCCTCCCACTGAAGTTTGTGCTTAAAA
AACAATAAATTTGACTTGGCACCATTGGGGTTGGTGGGAGAGGCCGTGTGACCTGGCTCTC
TGTCCAGTGCCACCAGGTATCCACATGCGCAG

FIGURE 228

MVNDRWKTMGGAAQLEDRPRDKPQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELLDTLADQLPRLARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRREDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAYELHVDLEDFENGTAARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSSLKHSGMRFTTKDRSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLRGAHASADG
VEWSSWTGWQYSLKFSEMKIRPVREDR

FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT
TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCT
CACCGCTGTAGGAATCCAG**ATG**CAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG
ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACCTCGGCATCCAGAGCCC
CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC
TTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTACTACC
AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCC
CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC
TGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG
AACAAATGGAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG
GACTGTAAATATTTCTGCCTTAGTGAAAACTCTACCATGCTGAAGATAAACAAACAAGAAGA
CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT
TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACCTCTGAACTG
TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG
GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA
TGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGACT**TGA**ATTCGCC
CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT
TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG
TTTCCTGTTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA
AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG
GCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTTCAT
GTCTTCCTTACACTTGGTGAATAAGAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC
ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC
AGCAAATACACAAGGAATCTTTTTGTTTGTTTCAGTTCATACTAGTCCCTTCCCAATCCAT
CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACCTGATATGAG
AATCTCAAATCTCAATGCCTTATAAGCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTG
TCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG
AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA
TCTGCACCTGTAATAGTTTCAGTTCCATTTTCTTCCATTGACCCATATTTATACCTTTCAG
GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

FIGURE 231

AATTTTCACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGNTGG
ATGATGATGGGACACCACCATGAGCCTGCATTNTCAAGCTTTTGCCACAATTCGGCATCCAG
AGCCCCGGCGCACAGAGCACAGGGNTCCTTTTTCAACGTGGCGACCAGTGGCCCTGACCCCTG
CTGACTTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTA
CTACCAGCTCTCCAATACTGGTCAAGACACCATTCTCAAATGGAAGAAAGATTAGGAAATA
CGTCCCAAGAGTTGCAATTTNTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTNTGCAGCAT
GTGGCTGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGGAACCTTGAAGGAGGGCAA
AGTNTCCTCATNTACTATACACACACCACTTCCC

FIGURE 232

GCCGAGCGCAAGAACCCTGCGCAGCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGGCTC
 CGGGGATTCGGCTCGGGCCGCTGGCTCTGCTCTGCGGGGAGGGAGCGGGCCCCGCCGCGGGG
 CCCGAGCCCTCCGGATCCGCCCCCTCCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT
 CTGGGGGTTCGCCGGGGCCGGGGACCCGCGGTCCGGGCGCC**ATG**CGGGCATCGCTGCTGCTG
 TCGGTGCTGCGGCCCCGAGGGCCCGTGGCCGTGGGCATCTCCCTGGGCTTCACCCTGAGCCT
 GCTCAGCGTCACCTGGGTGGAGGAGCCGTGCGGCCAGGCCCGCCCCAACCTGGAGACTCTG
 AGCTGCCGCCGCGCGGCAACACCAACGCGGCGCGCCGGCCCAACTCGGTGCAGCCCCGAGCG
 GAGCGCGAGAAGCCCCGGGGCCGGCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCTCTTGCC
 CTACCACCCTGCACAGCCCGGCCAGGCCGCCAAAAAGGCCGTGAGGACCCGCTACATCAGCA
 CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCCC
 ACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC
 GGGCGCACGGGGCCGCGGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC
 CCATTGGACACCTGCACCTGGCGCTGCGCCACCTGCTGGAGCAGCACGGCGACGACTTTGAC
 TGGTTCTTCTGGTGCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG
 CCACCTCAGCCTGGCCTCCGCCGCCACCTGTACCTGGGCCGGCCCCAGGACTTCATCGGCG
 GAGAGCCCCACCCCGGCCGCTACTGCCACGGAGGCTTTGGGGTGTCTGTGTCGCGCATGCTG
 CTGCAACAACCTGCGCCCCACCTGGAAGGCTGCCGAACGACATCGTCAGTGCGCGCCCTGA
 CGAGTGGCTGGGTGCTGCATTCTCGATGCCACCGGGGTGGGCTGCACTGGTGACCACGAGG
 GGGTGCCTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGCAGGAGGGGGACCCTCAT
 TTCCGAAGTGCCCTGACAGCCACCCCTGTGCGTGACCCTGTGCACATGTACCAGCTGCACAA
 AGCTTTTCGCGGAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA
 TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCGTGGGTATT
 CCAGCACCATCCCCGCCCGGCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA
 GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCCACTGCGTGGGGCTGACCGGGCTG
 ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCCTTG
 CGGCTCCAGAAGCAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCCGGGGTATGGA
 ATACACGCTGGACTTGCAGCTGGAGGCACTGACCCCCCAGGGAGGCCGCCGGCCCCCTCACTC
 GCCGAGTGCAGCTGCTCCGGCCGCTGAGCCGCGTGGAGATCTTGCCTGTGCCCTATGTCACT
 GAGGCCTCACGTCTCACTGTGCTGCTGCCCTTAGCTGCGGCTGAGCGTGACCTGGCCCCCTGG
 CTTCTTGGAGGCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC
 TGCTGCTACTGTATGAGCCGCGCCAGGCCAGCGCGTGGCCCATGCAGATGTCTTCGCACCT
 GTCAAGGCCCCAGTGGCAGAGCTGGAGCGGCGTTTCCCCGGTGCCCGGGTGCCATGGCTCAG
 TGTGCAGACAGCCGCACCCTCACCCTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC
 TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA
 CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCCAGAGCTGGGCCGTGACACTGGCCGCT
 TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACCTCCGACTACGTGGCAGCCCCGTGGG
 CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGAGCTGCTGGAGAGCCTGGATGTGTACGAGCT
 GTTCTCTCACTTCTCCAGTCTGCATGTGCTGCGGGCGGTGGAGCCGGCGCTGCTGCAGCGCT
 ACCGGGCCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC
 GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG
 CAACAGCACCT**TGA**CCCCACCCTGTCCCCGTGGGCCGTGGCATGGCCACACCCCAACCCACTT
 CTCCCCAAAACCAGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC
 AAGCTGGCCCACTGGTCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG
 GGAGTGCCCCCAGAGCCACCCACTTCTCATCCCAAACCCAGTTTCCCTGCCCCCTGACGCT
 GCTGATTGCGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC
 CTCTGGGCCCTGGGGGCTGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGGAGCTGAGGAGGGG
 GCATCTCCCAACTTCTCCCTTTTGGACCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA
 AGTGTGGAAAAA

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELP PRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAACKAVRTRYISTELGIRQRLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLLLE
QHGDDEFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCIL DATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFERSALTAHPVRDPVHMYQLHKAFARAELETTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGIPAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAAL TLLLLYEPRQAQRVA
HADVFAPVKAHVAELERR'PGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTL FLLAGPDTVL
TPDFLNRCRMHAISGWQAF FPMHFQAFHPGVAPPQGP GPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLES LDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQE QGNST

GCTCTGTGGCCGGCCCGCGATTGGTCAACGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT
TGGCAAGCGCTGGCCACCTCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT
AGCTATTAGCCAATTCTGGCAGGGGCCCGCTTTTTAGAAAGCTTGATTTCTTTGAAGATGAAAG
ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA
CTGTATCCACCCAAATGTCACCGATTTCTTCTCTATGCAGGAAATGAGCAGACCCATCAATAA
GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGCCAAAG
AGGGTTGCTCAACGCCCCGCCTCATTGAAAACCAAATCAGATCTGGGACCTATATAGCGTG
GCGGAGGCGGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT
TTCCCCGCCCTTGAGACCCTGCAGCACCATCTGTCT**ATG**GCGGCTGGGCTGTTTGGTTTGAGC
GCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCGGCCGCCCGCGTCCGCTGGGA
ATCTAGCTTCTCCAGGACTGTGGTGCCTCCGCTGTGGCGGGAAAGCGGCCCCCAAGAAC
CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACCTTGATGAGAAGAACCCA
GACTCCCATGGTTATGACAAGGACCCCGTTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT
CTTTGGCGTCTCCATCATCTCTGGTCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA
GGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC
CTTCCCATCATGGAATCCAACCTGCTTCGACCCAGCAAGATCCAGCTGCCAGAGGATGAG**TG**
ACCAGTTGCTAAGTGGGGCTCAAGAAGCACCGCCTTCCCCACCCCCTGCCTGCCATTCTGAC
CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

235/330

FIGURE 235

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE
DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER
LVKYREANGLPIMESNCFDPSKIQLPED

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 236

GGCGGCTGGGCTGTTTGGTTTGAGCGCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGC
TCCCGGCCGCCCCGCGTCCGCTGGGAATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCT
GTGGCGGGAAAGCGGCCCCCAGAACCGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGA
CGAAAACCTTGTATGAGAAGAACCCAGACTCCCATGGTTATGACAAGGACCCCGTTTTGGACG
TCTGGAACATGCGACTTGTCTTCTTCTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACC
TTTGTGGCCTATCTGCCTGACTACAGGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCT
TGTGAAATACCGAGAGGCCAATGGCCTTCCCATCATGGAATCCAACCTGCTTCGACCCCAGCA
AGATCCAG

CGCGCGGCT**ATG**CCGCGCTTGCTCTGCTCGTCTGTTGCTCCTGCGGGCCCGCGGCTGGTGCTT
TGCAGAACCCACGCGACAGCCTGCGGGAGGAACCTTGTTCATCACCCCGCTGCCTTCCGGGG
ACGTAGCCGCCACATTCCAGTTCCGCACGCGCTGGGATTTCGGAGCTTCAGCGGGAAGGAGTG
TCCCATTTACAGGCTCTTTCCCAAAGCCCTGGGGCAGCTGATCTCCAAGTATTCTCTACGGGA
GCTGCACCTGTCAATCACACAAGGCTTTTGGAGGACCCGATACTGGGGGCCACCCTTCCTGC
AGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCACTGATGTGGATAAA
TCTTGGAAGGAGCTCAGTAATGTCCTCTCAGGGATCTTCTGCGCCTCTCTCAACTTCATCGA
CTCCACCAACACAGTCACTCCCCTGCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTG
ACCACTACTTTCTGCGCTATGCTGTGCTGCCGCGGGAGGTGGTCTGCACCGAAAACCTCACC
CCCTGGAAGAAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTCTCTGTGCTGCTGAAGGCAGA
TCGCTTGTTCCACACCAGCTACCACTCCCAGGCAGTGATATCCGCCCTGTTTGCAGAAATG
CACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCTGTCAAGTTGTATTTGATGCCTTC
ATCACGGGGCAGGGAAAGAAAGACTGGTCCCTCTTCCGGATGTTCTCCCGAACCCCTCACGGA
GCCCTGCCCCCTGGCTTCAGAGAGCCGAGTCTATGTGGACATCACCACTACAACCAGGACA
ACGAGACATTAGAGGTGCACCCACCCCCGACCACTACATATCAGGACGTATCCTAGGCACT
CGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATCAACAACCTCTCGAAACCT
CAACATCCAGCTCAAGTGGAAGAGACCCCCAGAGAATGAGGCCCCCCCAGTGCCCTTCCTGC
ATGCCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTAC
AACACCCACCCATACCGGGCCCTTCCCGGTGCTGCTGCTGGACACCGTACCCTGGTATCTGCG
GCTGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAACAACCAAGTTACATCC
ACTACCAGCCTGCCCAGGACCGGCTGCAACCCACCTCCTGGAGATGCTGATTACAGCTGCCG
GCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTGAAGTGGACCGAGTA
CACGCCAGATCCTAACCATGGCTTCTATGTGACCCATCTGTCTCAGCGCCCTTGTGCCCA
GCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCCTCTTCAACAGCCTGTTCCCA
GTCTCTGATGGCTCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCC
GACACCGGACTTCAGCATGCCCTACAACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGT
GCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTCCACATCGAGGAGCCCCGCACAGGT
GGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCCCCTCT**TGA**TT
CTTGCCCTTTCCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTT
TCTGCCACTTGCTCTCCTCAGAGTTGGCTTTTGAACCAAAGTGCCCTGGACCAGGTACAGGC
CTACAGCTGTGTTGTCCAGTACAGGAGCCACGAGCCAAATGTGGCATTGGAATTTGAATTAA
CTTAGAAATTCATTTCTCACCTGTAGTGGCCACCTCTATATTGAGGTGCTCAATAAGCAAA
AGTGGTCGGTGGCTGCTGTATTGGACAGCACAGAAAAAGATTTCCATCACACAGAAAGGTC
GGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCACTGTGTAGTGA
TGGAGTTTACTGTTTGTGGAATAAAAAACGGCTGTTTCCGTGGAAAAA

FIGURE 238

MPLALLVLLLLGPGGWCLAEPDRSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPAKALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASFKPLGLANDTDHYFLRYAVLPREVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSSISWELRQTLVVFDAFITG
QGKKDWSLFRMFSRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRPPENEAPPVPFLHAQRYVSGYGLQKELSTLLYNTH
PYRAFPVLLLDTPWPYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLLVLNLPDPFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGIA
KRLANLIRRARGVPPL

FIGURE 239

CAAC**ATG**GGGTCCAGCAGCTTCTTGGTCCTCATGGTGTCTCTCGTTCTTGTGACCCTGGTGG
CTGTGGAAGGAGTTAAAGAGGGTATAGAGAAAGCAGGGGTTTGCCCAGCTGACAACGTACGC
TGCTTCAAGTCCGATCCTCCCCAGTGTACACAGACCAGGACTGTCTGGGGGAAAGGAAGTG
TTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCTGTGAAGGAACTGGAAGAAGGAGGAA
ACAAGGATGAAGATGTGTCAAGGCCATACCCTGAGCCAGGATGGGAGGCCAAGTGTCCAGGC
TCCTCCTCTACCAGGTGTCCTCAGAAA**TGA**TGCTGGGTCTTTCTACCTCTGGGGGTCACTC
TCACTTGGCACCTGCCCCTGAGGGTCCTGAGACTTGGAATATGGAAGAAGCAATACCCAACC
CCACCAAAGAAAACCTGAGCTTGAAGTCCTTTTCCCCAAAAAGAGGGAAGAGTCACAAAAG
TCCAGACCCCAGGGACGGTACTTTCCCTCTCTACCTGGTGCTCCTCCCTAATGCTCATGAAT
GGACCCCTCATGAATGAAACCAGTGCCCTTATAAGAGACCCCAAAGAGCTGCCTTGCCCTTC
TGCAATGTGTGATCACAGCTAGAAGGCACTGTCAGAGAAGAGAACTGGTCCTCACCAGATG
CTGAATCTGCTGGTGCCTTGATCTTGGACTTCCCAGCCTCTAGAACTGTAAGAAATAAATAT
TTGCTGTTTATAATCCAA

240/330

FIGURE 240

MGSSSFLVLMVSLVLTSLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC
YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

Signal sequence:

amino acids 1-19

N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

240/330

AAACCTCAGCACTTGCCTGGAGTGGCTCATTTGTTAAGACAAAGGGTGTGCACCTTCTCTGGCCAGG
AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC
TCTAGAACCCGACCCACCACC**ATG**AGGTCCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGG
CGTCCAGTGGTCCTTGCTTCTGGCTGTCTGCTGCTTCTTTCTCTTGCCTTGGCCTCTTTTA
TTAAGGAGCCTCAAACAAAGCCTTCCAGGCATCAACGCACAGAGAACATTAAAGAAAGGTCT
CTACAGTCCCTGGCAAAGCCTAAGTCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTA
TGCAGAGCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCCAGCCCAAGGCCACACCA
CCGGAGACAGAGGAAAGGAGGCCAACCAGGCACCGCCGGAGGAGCAGGACAAGGTGCCCCAC
ACAGCACAGAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAAACCATGGTGAACACACTGTC
ACCCAGAGGGCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCC
AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG
GTGTCTAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTTCCAAAAGTCA
GCACAGAATGCTGGCTCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA
CAGCAGTCATCCACCTAAGGAGAAGAAACCTCAGGCCACCCACCCCTGCCCTTTCCAG
AGCCCCACGACGCAGAGAAACCAAGACTGAAGGCCGCAACTTCAAATCTGAGCCTCGGTG
GGATTTTGGAGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG
TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAAACTCTTTCTGCCAACCTCACTCTC
TTCTTGACTCCAGACACTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACC
CTTTGGCTTCATGGAGCTCAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCCTCCAG
TGCCCCAGCAGCAGCTGCTCCTGGCCAGCCTCCCGCTGGGAGCCTCCGGTGCATCACCTGT
GCCGTGGTGGGCAACGGGGGCATCCTGAACAACTCCACATGGGCCAGGAGATAGACAGTCA
CGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAAGGCTACGAACAGGATGTGGGGACTC
GGACATCCTTCTACGGCTTTACCGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAAT
CGGGGTTTTCAAGAACGTGCCTCTTGGGAAGGACGTCCGCTACTTGCACTTCTTGGAAGGCAC
CCGGGACTATGAGTGGCTGGAAGCACTGCTTATGAATCAGACGCTGATGTCAAAAAACCTTT
TCTGGTTTCAGGCACAGACCCAGGAAGCTTTTTCGGGAAGCCCTGCACATGGACAGGTACCTG
TTGCTGCACCCAGACTTTTCTCCGATACATGAAGAACAGGTTTTCTGAGGTCTAAGACCCTGGA
TGGTGCCCACTGGAGGATATACCGCCCCACCACTGGGGCCCTCCTGCTGCTCACTGCCCTTC
AGCTCTGTGACCAGGTGAGTGCTTATGGCTTCATCACTGAGGGCCATGAGCGCTTTTCTGAT
CACTACTATGATACATCATGGAAGCGGCTGATCTTTTACATAAACCATGACTTCAAGCTGGA
GAGAGAAGTCTGGAAGCGGCTACACGATGAAGGGATAATCCGGCTGTACCAGCGTCCTGGTC
CCGGAACCTGCCAAAGCCAAGAACT**TGA**CCGGGGCCAGGGCTGCCATGGTCTCCTTGCCCTGCTC
CAAGGCACAGGATACAGTGGAATCTTGAGACTCTTTGGCCATTTCCCATGGCTCAGACTAA
GCTCCAAGCCCTTCAGGAGTTCGAAGGGAACACTTGAACCATGGACAAGACTCTCTCAAGAT
GGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCAGTACATTGCTGTAGGTCCTGAGGCCAGG
GATTTTTTAATTAAATGGGGTGATGGGTGGCCAATACCACAATTCCTGCTGAAAAACACTCTT
CCAGTCCAAAAGCTTCTTGATACAGAAAAAAGAGCCTGGATTTACAGAAACATATAGATCTG
GTTTGAATTCCAGATCGAGTTTACAGTTGTGAAATCTTGAAGGTATTACTTAACTTCACTAC
AGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTCTAGAAGGGTCTATACTTGTCCCTTG
TCTTTAAGCTATTTGACAACTCTACGTGTTGTAGAAAACTGATAATAATACAAATGATTGTT
GTCCATGGAAAGGC AAATAAATTTTCTACAGTGAAAAAAAAAAAAAAAAA

FIGURE 242

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQNGGGQTRKLTASRTVSEKHQG
KAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPEKKPQATPPAPFQSPTTQRN
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFPLNLTLFLDSRHF
NQSEWDRLEHFAPPFGFMELNYSLVQKVVTFRFPVPQQQLLLASLPAGSLRCITCAVVGNGG
ILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFATAFSLTQSLLILGNRGFKNVP
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
RYMKNRFLRSKTLGDGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDDHYDTSW
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

FIGURE 243

CG**ATG**CGCGGACCCGGGCACCCCCTCCTCCTGGGGCTGCTGCTGGTCTGGGGCCTTCGCCG
GAGCAGCGAGTGGAAATTGTTCTCGAGATCTGAGGATGAAGGACAAGTTTCTAAAACACCT
TACAGGCCCTCTTTATTTTAGTCCAAAGTGCAGCAAACACTTCCATAGACTTTATCACAACA
CCAGAGACTGCACCATTCCTGCATACTATAAAAGATGCGCCAGGCTTCTTACCCGGCTGGCT
GTCAGTCCAGTGTGCATGGAGGATAAG**TGA**GCAGACCGTACAGGAGCAGCACACCAGGAGCC
ATGAGAAGTGCCTTGGAACCAACAGGGAAACAGAACTATCTTTATACACATCCCCTCATGG
ACAAGAGATTTATTTTTGCAGACAGACTCTTCCATAAGTCCTTTGAGTTTTGTATGTTGTTG
ACAGTTTGCAGATATATATTCGATAAATCAGTGTACTTGACAGTGTTATCTGTCACTTATTT

244/330

FIGURE 244

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFRLYHNT
RDCTIPAYYKRCARLLTRLAVSPVCMEDK

244/330

FIGURE 245

GGGCTGGGCCCCGCCGAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG
CCCGACCCCGGCCGCGCCAGCCCCACC**ATG**CCACCCGCGGGGCTCCGCCGGGCCGCGCCG
CTCACCGBAATCGCTCTGTTGGTGCTGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT
GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT
GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCAGAGAGG
CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT
CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCCTCTGTTCCCTGTTGCTACCTGT
ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTGAAGGCCAGGAGATTCCAATGACAGGCATC
CCAGTGACGCCAGTATACCCATACCCCCAGGACCCCAAAGCTGGCCCTGCACCCCCACAGCC
TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC
CAGTCTACAACCCCTGCAGCTCCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC
TGAGGGAACCAGCCATGTCTCTGCTGCCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT
CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA
GCCAAGCCCTGGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA
ACTATGAGGGGTGGGGGGAGGGCTTGGAATTATGGGCTATTTTACTGGGGGCAAGGGAGG
GAGATGACAGCCTGGGTACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG
CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGGAGGAGGT
TCCGTCAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG
CTAGATTAAAGCTGTAAAGACAAAA

FIGURE 246

MPPAGLRRRAAPLTAIALLVLGAPLVLAGEDCLWYLDNRNGSWHPGFNCEFFTFCGTCYHRYC
CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCFLCSCCYLYRRRQQLQSP
FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPPP
YMPPQPSYPGA

Transmembrane Domains:

amino acids 10-28, 85-110

N-glycosylation Site:

amino acids 38-41

N-myristoylation Sites:

amino acids 5-10, 88-93

246/330

FIGURE 248

MVFLPLKWSLATMSFLLSLLALLTVSTPSWCQSTEASPKRSDGTFPWNKIRLPEYVIPVH
 YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLOVLE
 HPPQEQIALLAPEPLLVLGPLYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA
 ARMAFPCFDEPAFKASF SIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA
 FIISDFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAA
 IPDFQSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL
 WLNFGFAKFMEFVSVSVTHPELVGVDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD
 DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDG
 FCSRSQHSSSSSHWHQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD
 TGYLWHVPLTFITSKSNMVHRFLLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL
 TGLLKGTHTAVSSNDRASLINNAFQLV SIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIP
 MYKLMEKRD MNVETQFKAF LIRLLRDLIDKQTTWDEGSVSEQMLRSELLLLACVHNYQPCV
 QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSSTEKSQIEFALC
 RTQNKEKLQWLLDESFKGDKIKTQEFQILTLIGRNPVGYPLAWQFLRKNWNKLVQKFELGS
 SSIAHVMGT TNQFSTRTRLEEVKGFFSSLKENGSQLRCVQQTIETIEENIGWMDKNFDKIR
 VWLQSEKLERM

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

248/330

FIGURE 249

CAGCCACAGACGGGTC**ATG**AGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCAC
TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC
GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA
CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG
AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC
TCCTACACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCT
TTGGGCCCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCCAGTCTGCTTGTCTATGG
AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT
GGCCTCCTCAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCC
CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCCGTGGGTATGACTGAGA
ACTGCAATAGGAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTTATGACACACGGAAAC
TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT
GTGTCAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG
GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCCTCCTGGG
GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG
CAGCGTTCTGCTGAACTCCCTCCCTCCTCAAGCTGCCCCCTGTCCCAGGAGACCGGCAGTGTC
CTACCTGTGTGCAGCCCCCTTGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG
GGCGCCACTCATTGTTATGATGGGTACATTATCTCTCAGGAGGTGGGCTGTCCACCAAAT
GAGCATTGAGGGCTGCGTGGCCCCAACCTTCCAGCTTCTTGTGAACCACACCAGACAAATCG
GGATCTTCTCTGCGCGTGAGAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGT
GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG
GTGGGGAGTGGTTTGCCCTTCCTGCT**TAA**CTCTATTACCCCCACGATTCTTCACCGCTGCTGA
CCACCCCACTCAACCTCCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC
CCATTCTGTCCATGAATCATCTTCCCCACACACAATCATTCATATCTACTCACCTAACAGCA
AACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG
GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

FIGURE 250

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVKVSDLPRQWTPKNTSCDSGLGCQDTLMLI
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTTHCYDGLLRRLRGGGIFSNLRVQGCMPQPGCN
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSENTEMCEVGQVCQETL
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESITWGVGLALAPALWWGVVC
PSC

P0511 23.06.00

FIGURE 252

MRGNLALVGVLI SLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGRVG
PTGEKGMGDKGQKGSVGRHGKIGPIGSKGEKGDSDIGPPGPNGEPGLPCECSQLRKAIGE
MDNQVSQLTSELKFIKNAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMPKDEAAN
GLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPPNNAYDEEDCVEMVAS
GGWNDVACHTTMYFMCEFDKENM

252/330

FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG
CACCAGTGTGTGAGGGGAGCAGGCAGCGGTCCTAGCCAGTTCCTTGATCCTGCCAGACCACC
CAGCCCCCGGCACAGAGCTGCTCCACAGGCACCATGAGGATCATGCTGCTATTACAGCCAT
CCTGGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGG
TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC
AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA
GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA
GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCCTCGGCCC
CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC
TTTATTAAGACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT
ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC
TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTCCCTGTTCTTGT
AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA
TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT
CCTACATTAAAAATATAATGTCTCTCTCTATTCTCAACAATAAAGGATTTTTGCATATGAA
AA

FIGURE 254

MRIMLLFTAILAFSLAQSFQAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK
ALSQASTDPKESTSPEKRDMHDFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK
SSLGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCGCTGTGCCTGTGTGCC
CGCGCTGTCGCCGCTGCTACCGCTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG
GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC
CCATAGTTGCTGCAGGAGTGGAGCCATGAGCTGCGTCCTGGGTGGTGTTCATCCCCTTGGGGC
TGCTGTTCTGGTCTGCGGATCCCCAAGGCTACCTCCTGCCCAACGTCACCTCTCTTAGAGGAG
CTGCTCAGCAAAATACCAGCACAAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA
GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGGCCAGGTGCAGCCTCAGGCCT
CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC
CTGGGGCCACCAGCCTGCTCTGTTCCCCAGCCAGCTCTGTTCCCCAGCCAGTGCGTGTGATGG
CTGGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCGGCTCTGTTCTGTTTTGTTTGTGTTGTT
TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG
AAACCTTAGACTCCCGGGGTAAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACCTACAG
GCATGCACCATGGTGCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT
TGCCCAGGCTGGTCTTGAACCTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG
CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA
ACAACACACGTGGGTTCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC
TCCACTGGGAACACAGCTCTCAGCCTTTCCACCTGGAGGCAGAGTGGGGAGGGGGCCCAGGG
CTGGGCTTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC
CTTAGCCCGTGTTGAGCCTCACTTTCCACTTGGAGAGTCCTTTCCTCGCGTGGTTGCCATGACT
GTGAGATAAGTCGAGGCTGTGAAGGGCCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG
CTTTGCTAACCAGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT
GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG
TGGAAACTTCCTTCCTGGCTGGTTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG
GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT
TTTAAACAAACCCACGTCCCAGCCTGGGTAACATGGTAAAGCCCCGTCTCTACAAAAAAATC
CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG
GTGGAGGTGGGGGGTGGGAGCTGAGGAAGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC
AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

FIGURE 256

MSCVLGGVIPLGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNEHSRVRRAIPREDKEEILML
HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR
GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

FIGURE 257

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACT**ATGGG**
GTCTGGGCTGCCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGG
GTATGACTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGC
TTCCTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCAC
CCTCCACCATGCAAGATCTCAACACCATGTTGTCTGCAACACA**TGA**CAGCCATTGAAGCCTG
TGTCTTCTTGGCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTT
CAGCAGGCCCCCACCCTCCTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 259

AATTGTATCTGTGTAATGTTAAAACAAACGAAATAAAATAGAAGGAAAACTTTCTGAGTTT
CAAAAACAACAGACTAGTACTCTAAAGAACTCTTTAAAACAATTAAGTGTAGGATTGCAGT
T**ATG**ATTGGATATTATTTAATTCTGTTTCTGATGTGGGGTTCCTCCACTGTGTTCTGTGTGC
TATTAATATTTACCATTGCAGAAGCTTCATTCAGTGTTGAAAATGAATGCTTAGTGATCTG
TGCCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAGTTCCCCTCCC
CTCCGATTGTTCTAAAT**TAA**TTGAAAGATGTCTGCTGTGGAAAAAGGCATGTATTTAAATCTG
TATGATTCTCAACCATCTTTAGTTGGGAAAGGTCCCTTGAAAGCCAATGGAAATACTTTTTTTT
TTTTCTTGGCACTAATCAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTA
GTAAAATAGAAACCTGTGTTTATTCTCAGGTATTTTAGAAACAACAGCCATCATTTTATTTT
ATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGGGCTATCAAATATTACTTCAT
TCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTTCTAGTTGCATTTTCTC
AGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTTACAGTACTGTCTCTCTACTA
TCTTCAGATTACTTGATTCAAATAAACCAATTATGTTTGTAATTGATATTAATAAAACCAGA
ATAAAAGTTCATATCTACCC

260/330

FIGURE 260

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFVENECLVDLCLLRICYKLSGVPNQCRVPLP
SDCSK

Important features:

Signal peptide:

amino acids 1-29

260/330

FIGURE 261

GAGGATTTGCCACAGCAGCGGATAGAGCAGGAGAGCACCACCGGAGCCCTTGAGACATCCTT
 GAGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTGTTTTGCAGG**ATG**ATGGTGGCCCTT
 CGAGGAGCTTCTGCATTGCTGGTTCTGTTCCTTGACAGCTTTTCTGCCCCCGCCGAGTGTAC
 CCAGGACCCAGCCATGGTGCATTACATCTACCAGCGCTTTCGAGTCTTGGAGCAAGGGCTGG
 AAAAATGTACCCAAGCAACGAGGGCATAACATTCAAGAATTCGAAGAGTTCTCAAAAAATATA
 TCTGTCATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCAGTGGGTAACCTT
 GGCAGTGCAGAGTTGAACGTGCCCAACGGGAGATTGACTACATAACAATACCTTCGAGAGGCTG
 ACGAGTGCATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAA
 GAGAAAAAGATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTC
 TTTGAAAATAGTGAAGAAGATGATGGACACACATGGCTCTTGGATGAAAGATGCTGTCTATA
 ACTCTCCAAAGGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAAC
 ATACGGGCATTTCATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAAATCCTAACACTTTC
 CTGGCAGGGGAACAGGCCAAGTGATCTACAAAGGTTTTCTATTTTTTTCATAACCAAGCAACTT
 CTAATGAGATAATCAAATATAACCTGCAGAAGAGGACTGTGGAAGATCGAATGCTGCTCCCA
 GGAGGGGTAGGCCGAGCATTTGGTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGT
 GGATGAGCATGGGCTCTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCA
 CAAAGATTGAGCCGGGCACACTGGGAGTGGAGCATTCATGGGATACCCCATGCAGAAGCCAG
 GATGCTGAAGCCTCATTCTCTTGTGTGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCA
 GGGCCCTCATCGCATCACCTGCATCTATGATCCACTGGGCACTATCAGTGAGGAGGACTTGC
 CCAACTTGTTCTTCCCCAAGAGACCAAGAAGTCACTCCATGATCCATTACAACCCAGAGAT
 AAGCAGCTCTATGCCTGGAATGAAGGAAACCAGATCATTTACAAACTCCAGACAAAGAGAAA
 GCTGCCTCTGAAG**TAA**TGCATTACAGCTGTGAGAAAGAGCACTGTGGCTTTGGCAGCTGTTT
 TACAGGACAGTGAGGCTATAGCCCCTTCAACAATATAGTATCCCTCTAATCACACACAGGAAG
 AGTGTGTAGAAGTGGAATACGTATGCCTCCTTTCCCAAATGTCACTGCCTTAGGTATCTTC
 CAAGAGCTTAGATGAGAGCATATCATCAGGAAAGTTTCAACAATGTCCATTACTCCCCAAA
 CCTCCTGGCTCTCAAGGATGACCACATTCTGATACAGCCTACTTCAAGCCTTTTGTCTTACT
 GCTCCCCAGCATTTACTGTAACTCTGCCATCTTCCCTCCCACAATTAGAGTTGTATGCCAGC
 CCCTAATATTACCACTGGCTTTTCTCTCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTT
 CAAATGTCTATTGATATTCTCCATTTTCACTGCCCACTAAAATACTATTAATATTTCTTT
 CTTTTCTTTCTTTTTTTTTGAGACAAGGTCTCACTATGTTGCCAGGCTGGTCTCAAACCTCC
 AGAGCTCAAGAGATCCTCCTGCCTCAGCCTCCTAAGTACCTGGGATTACAGGCATGTGCCAC
 CACACCTGGCTTAAAATACTATTTCTTATTGAGGTTTAACTCTATTTCCCTAGCCCTGTC
 CTTCCACTAAGCTTGGTAGATGTAATAATAAAGTGAAAATATTAACATTTGAATATCGCTTT
 CCAGGTGTGGAGTGTGTCACATCATTGAATTCTCGTTTACCTTTGTGAAACATGCACAAG
 TCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGTAACACAATTACAAAGTGAAAGATACAGC
 TAGAAAATACTACAAATCCCATAGTTTTTCCATTGCCCAAGGAAGCATCAAATACGTATGTT
 TGTTACCTACTCTTATAGTCAATGCGTTTCATCGTTTACGCCTAAAAATAATAGTCTGTCCC
 TTTAGCCAGTTTTTCATGTCTGCACAAGACCTTTCAATAGGCCTTTCAAATGATAATTCCTCC
 AGAAAACCAGTCTAAGGGTGAGGACCCCACTCTAGCCTCCTCTTGTCTTGCTGTCTCTGT
 TTCTCTCTTTCTGCTTTAAATTCAATAAAAGTGACACTGAGCAAAAAAAAAAAAAA

FIGURE 262

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQE
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL
QEAEKKKIRTLNLASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNPKVYLLIGSRNNTV
WEFANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKEPGLGVEHSWDT
PCRSQDAEASFLLCGVLYVVYSTGGQGPFRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH
YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQIRIRKEKPQQHNFTHRLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTHAAVSPCGRFVASC
FTPDKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFVRH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ
AQETLKSLGALKK

Important features:**Signal peptide:**

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

FIGURE 265

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGG
 CAGTGTTTTGCCTTCACCCCAAGTGACC**ATG**AGAGGTGCCACGCGAGTCTCAATCATGCTCC
 TCCTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCCCTGTGAGCGGGATGTCCAGTGT
 GGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT
 GGGGCGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA
 AGCACCACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTAC
 CGCTGCTCCATGGACTTGAAGAACATCAATTTT**TAG**GCGCTTGCCTGGTCTCAGGATACCCA
 CCATCCTTTTCCTGAGCACAGCCTGGATTTTTATTTCTGCCATGAAACCCAGCTCCCATGAC
 TCTCCCAGTCCCTACACTGACTACCCTGATCTCTCTTGTCTAGTACGCACATATGCACACAG
 GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTACAGCTTGAGG
 CTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTA
 AATGGCAGAAAGGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCCCTGGGCCCTG
 CCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCCTGGGCACAGGCTCTTGGGT
 GCATTGCTCAGAGTCCCAGGTCCTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG
 ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC
 TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA
 GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA
 CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGT
 TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAATGCAAATGGTGGTGGTTCAA
 TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACCTCTTTCCA
 AGGATCAGCCCTGAGAGCAGGTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG
 TGGGAGCAAGGGACAGGGAGCAGGGCAGGGGCTGAAAGGGGCACTGATTTCAGACCAGGGAGG
 CAACTACACACCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAA

FIGURE 266

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEECHP
GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

FIGURE 267

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACCTTGAGAAATGTCTTTC
CTCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGC
TGCTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGG
CCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAA
GCAAAGGAGCTATGGGAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTT
CCTCTGTGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCG
TCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTAT
TTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAAGTTCTATGGTCCACAAAGGCGGAAGAT
GATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACCTCTTCCGAGCCTGGAACGGAG
GCTTCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTCTGTTGGGATCA
GGAAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACT
TTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAATGAT
TGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTACCTGTGTTTCATGGGATGTATT
GTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTA
TTAATGTATTTTAATATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGA
CAAAAATCTGAAAACTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAA
TTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG
TGAGCAAGTCACTTGAGGTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTC
TCTACTAAAAATACAAAATCACCCGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCG
GGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCA
CACCCTGTATTCCAGCCTGGGTGACTGAGACTCTAACTAA

MSFLQDPSFFTMMWSIGAGALGAAALALLLANTDVFLSKPQKALEYLEDIDLKTLEKEPR
TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF
QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGVWYNFFRAWNGGFSGNLEGEFGFILGGVFV
VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

FIGURE 269

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCG
GGCCAGGTGCCCCGTCGCAGGTGCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGA
AGCCCCTTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTG
CTTCTGGCGCTGGGCCTGCCGTTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACA
GACCACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATG
GCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTG
CTCCTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGG
CACCTACCGGCCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGG
ACTCCAAGGAGACGGTGCAGGGCTGCCTGCCCATCTAGGTCCCCTCTCCTGCATCTGTCTCC
CTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAG
TGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCTGAGGTCAAGAGAGGATGGG
GCTATTCACCTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 270

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIV
VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

FIGURE 271

AATATATCATCTATTTATCATTAAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTT
TTGGGATTTTAAATTTTCAAACACAGCAGAAATGACATTTTTTCTGTCACTATTATTATTGTTG
GTATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCT
ATCAAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTC
CTACCAAAGCTGTCAAACCACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGA
GGGTTAATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACTTAGATTCAATG
ATTGTAAATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATA
TATACAATAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACT
TTATTAATTTTTTAAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGA
TCATATAATTTGATACAAATAAAAGAAAAGTGTTCTCTCCCCTTACAGAATTGACATTTTAA
ATGCGATACAGTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAA
AGAAGGGAAAATGTTGCCAAGGAAAAAAAAA

272/330

FIGURE 272

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK
GIVKGRNLDSRGLILGAEAWGRGVKKNT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 273

GCCAGGAATAACTAGAGAGGAACA**ATG**GGGGTTATTTCAGAGGTTTTGTTTTCTCTTAGTTCT
 GTGCCTGCTGCACCAGTCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA
 TTGTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAATAATTGAACAAATAGAGGAT
 ATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGATTTTTTTTCAAAAA
 TGTATCTATATTAATTCCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATG
 AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA
 TACACCAAGCAGTTCACAGAATGTGGAGAGAAAGGCCAATACATTCACCTTCACCCCTGACCT
 TCTACTTGGAAAAAAACAAAATGAATATGGACCACCAGGCCAACTGTTTGTCCATGAGTGGG
 CTCACCTCCGGTGGGGAGTGTGTTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG
 TCAAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA
 GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACCTGTATG
 GAAAAGATTGTCAATTCTTTCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG
 CAAAGTATTGATTCTGTTGTTGAATTTTGTAAACGAAAAAACCCATAATCAAGAAGCTCCAAG
 CCTACAAAACATAAAGTGCAATTTTGAAGATCAGGAGGTGATTAGCAATTCCTGAGGATT
 TTA AAAACACCATAACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTCGTGAAGATC
 AGTCAAAGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCT
 AAATCGAATGAATCAAGCAGCAAAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGG
 TGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTAAATAAGCTAATCCAAATAAAAAGC
 AGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACCTCCAT
 CTGCTCTGGAATTAATATGCATTTTCAAGGTGATTGGAGAGCTACATTTCCCAACTCGATGGAT
 CCGAAGTACTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTG
 AAACAAAGTGGGGCCATTGTTTCATTTTATTTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT
 AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAATG
 GCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAGTCCCTT
 CAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT
 TGATAGTACAGTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCAGTA
 TTTCTCTCTGGGATCCAGTGGAACAATAATGGAATAATTTTACAGTGGATGCAACTTCCAAA
 ATGGCCTATCTCAGTATTCCAGGAACCTGCAAAGGTGGGCACCTTGGGCATACAATCTTCAAGC
 CAAAGCGAACCAGAAACATTAACATTACAGTAACCTTCTCGAGCAGCAAATCTTCTGTGC
 CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTTCCCGAGCCCAATGATT
 GTTTACGCAGAAATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTTCAT
 TGAATCACAGAATGGACATACAGAAGTTTTTGGAACTTTTGGATAATGGTGCAGGCGCTGATT
 CTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAATGGCAGATAT
 AGCTTAAAAGTTCGGGCTCATGGAGGAGCAAAACACTGCCAGGCTAAAATTACGGCCTCCACT
 GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAACCCGCCAA
 GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTTCAAGCCGAACAGCATCCGGA
 GGTGCATTTGTGGTATCACAAGTCCCAAGCCTTCCCTTGCTTGACCAATACCCACCAAGTCA
 AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTTACATGGACAGCACCAG
 GAGATAATTTTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT
 GATCTAAGAGACAGTTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCAACAAAGGA
 GGCCAACCTCCAAGGAAAGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCC
 ACATATTTATTGCCATTAAAAGTATAGATAAAAGCAATTTGACATCAAAGTATCCAACATT
 GCACAAGTAACCTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC
 TACTCCTACTCCTACTCCTGATAAAAGTCATAATTTCTGGAGTTAATATTTCTACGCTGGTAT
 TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATT**TGA**ACCTTA
 ACGAAGAAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTTAAAAACAAAACAATGTAAGT
 AAAGGATATTTCTGAATCTTAAAATTCATCCCATGTGTGATCATAAACTCATAAAAATAATT
 TTAAGATGTCGAAAAGGATACTTTGATTAAATAAAAACACTCATGGATATGTA AAAACTGT
 CAAGATTA AAAATTTAATAGTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGAAAC
 AAAGATCCTTTTTTCACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAAT
 GATATTTCAAATTCATCAAGAAATTA AAATCATCTATCTGAGTAGTCAAAATACAAGTAAA
 GGAGAGCAAATAAACACATTTGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 274

MGLFRGFVFLLVLCLLHQSNTSFIKLNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTY
 LFEATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTEC
 GEKGEYIHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR
 CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVVE
 FCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPVFSLKISQRIVCLV
 LDKSGSMGGKDRLNRMNQAAKHFLQLTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM
 AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTGDGDNNTASSCIDEVKQSGAIVH
 FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
 LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSTIMENFTVDATSKMAYLSIPG
 TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG
 YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNQGVYSRYFTAYTENGRYSLKVRAGH
 GANTARLKLRPPLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRASGGAFVVSQV
 PSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDNFQVGVQRYIIRISASILDLRDSFDD
 ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIAQVTLFIP
 QANPDDIDPTPTPTPTPTPTDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

[illegible]

FIGURE 276

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIES
ALETLNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSP
MKVKNEKLNMAN TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVK
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLN TCR
LNTQPGKKVFYPVLF SQYNPGIIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFI
NIGGFDL DIKGWGGEDVHLYRKYLHSNLIVVRTPV RGLFHLWHEKRCMDELTP EQYKMCMQS
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

FIGURE 277

GAAAGA**ATG**CTTTGTTGGCTGCTCTTTTTTCTGGTGACTGCCATTTCATGCTGAACTCTGTCAACC
AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT
ATGCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA
GTTCCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT
ATCATTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC
AATCAGCCATAAGAATGAACAAGAACCGGATCAACAATGCCTTCTTTCTAAATGACCAA
CTGGAATTTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTG
GATTATTATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTAT
CAGGGATCTGGCAACGTAGAAGAAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT
AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCTCTGATCCCCTGGACATGAAGGG
GGGCATATTAATGATGCCTTCA**TGA**CAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGT
TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA
CCAAGAGCAGATCATATATTTTGTTCACCATTCTTCTTTTGAATAAATTTTGAATGTGCT
TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGAC
TCAAATATTCTAAAATATTTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG
TAGTTATTGATTTAAGCATTTTTTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTC
AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT
TGAAAATGGATCCTTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG
TGAGAAGTAATTATTGTAAATGGATGGATAAAAAATGGAATTACTCATATACAGGGTGGAAATT
TTATCCTGTTATCACACCAACAGTTGATTATATATTTTTCTGAATATCAGCCCCTAATAGGAC
AATTCTATTTGTTGACCATTTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG
TAATAATCATCTCTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

278/330

FIGURE 278

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVAFSMRKVP
NREATEISHVLLCNVTQRVSFVFWVTDPSKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLE
FLKIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLLILSGIWQRRRKNKEPSEVDDAEDKC
ENMITIENGIPSDPLDMKGGILMMPS

278/330

FIGURE 279

AACTCAAACCTCTCTCTGCGGAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG
 GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTCATGT
 ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGG
 CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC
 TCGGTTAAATGCACTTTCTCCAGCTTTGCCCCGTGTGGGTGATGCTCTAACAGTGACCTGGA
 ATTTTCGTCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC
 CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA
 TGCCTCCATCCTTCTCTGGAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA
 AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA
 CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT
 AATAGTAATTGTAGTGGTCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC
 ATAAAGTGGTGGAGATAAAATCAAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT
 GTTTATTTAGAAGACACAGACTAAACAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA
 GAACCCTAGTATTTCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT
 TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC
 AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA
 GTGTAAATTTTTTCAAGTGCTCATTAGGTTTATAAACAAGAAGCTACATTTTTTGCCCTTAA
 GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC
 AATTTGTCTGTTACATTTCTTTTACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA
 ATGTGTTTACTCTCTTTCTTCCACATTCTCAATTAAAGGTGAGCTAAGCCTCCTCGGTG
 TTTCTGATTAACAGTAAATCCTAAATTCAAACGTGTTAAATGACATTTTTTATTTTTATGTCTC
 TCCTTAACTATGAGACACATCTTGTTTTACTGAATTTCTTTCAATATTCAGGTGATAGATT
 TTTGTCTG

FIGURE 280

MYGKSSTRAVLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVT
WNFRPLDGGPEQFVFYYHIDPFQPMSEGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ
VKNPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIQSACALMIIIVIVVVLFFQHYRKKRWAER
AHKVVEIKSKEEERLNQEKKVSVYLEDTD

280/330

FIGURE 281

GCATTTTTGTCTGTGCTCCCTGATCTTCAGGTCACCACC**ATG**AAGTTCTTAGCAGTCCTGGT
ACTCTTGGGAGTTTCCATCTTTCTGGTCTCTGCCCAGAATCCGACAACAGCTGCTCCAGCTG
ACACGTATCCAGCTACTGGTCCTGCTGATGATGAAGCCCCTGATGCTGAAACCACTGCTGCT
GCAACCACTGCGACCACTGCTGCTCCTACCACTGCAACCACCGCTGCTTCTACCACTGCTCG
TAAAGACATTCCAGTTTTACCCAAATGGGTTGGGGATCTCCCGAATGGTAGAGTGTGTCCCT**T**
GAATGGAATCAGCTTGAGTCTTCTGCAATTGGTCACAACCTATTCATGCTTCCTGTGATTTT
ATCCAACCTACTTACCTTGCTACGATATCCCCCTTATCTCTAATCAGTTTATTTTCTTTCAA
ATAAAAAATAACTATGAGCAACATAAAAAAAAAAAAAA

282/330

FIGURE 282

MKFLAVLVLLGVSI FLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTAT
TAASTTARKDIPVLPKWVGDLPNGRVCP

282/330

FIGURE 283

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC
CTAGGGGTCTGGATTTGCTGGTTAACAAGATAACCTGAGGGCAGGACCCCATAGGGGA**ATGC**
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTACGGTGGCCTGGTCCCTCCTTGCCGAGAGA
GTGTCCTGGGTCAGGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGGAC
ACTTGGCAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGGG
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGGCCAGTCCAGGGTGGGGGGCG
GCAAACCTCCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCCAGCTCCT
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGCTTTGTGCCAG**TAAA**ACCACAGGCTGG
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTTTACCACATTAGCAATTAAAACCTGAGAAAT
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA
TACAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCGCCGCTGAT
TCCAGCCTGGGCGACAAGAGTGAGACTCCATCTCACACA

FIGURE 285

GTCATGCCAGTGCCTGCTCTGTGCCTGCTCTGGGCCCTGGCAATGGTGACCCGGCCTGCCTCA
GCGGCCCCCATGGGCGGCCCAGAACTGGCACAGCATGAGGAGCTGACCCTGCTCTTCCATGG
GACCCTGCAGCTGGGCCAGGCCCTCAACGGTGTGTACAGGACCACGGAGGGACGGCTGACAA
AGGCCAGGAACAGCCTGGGTCTCTATGGCCGCACAATAGAACTCCTGGGGCAGGAGGTCAGC
CGGGGCCGGGATGCAGCCCAGGAACTTCGGGCAAGCCTGTTGGAGACTCAGATGGAGGAGGA
TATTCTGCAGCTGCAGGCAGAGGCCACAGCTGAGGTGCTGGGGGAGGTGGCCCAGGCACAGA
AGGTGCTACGGGACAGCGTGCAGCGGCTAGAAGTCCAGCTGAGGAGCGCCTGGCTGGGCCCT
GCCTACCGAGAATTTGAGGTCTTAAAGGCTCACGCTGACAAGCAGAGCCACATCCTATGGGC
CCTCACAGGCCACGTGCAGCGGCAGAGGCGGGAGATGGTGGCACAGCAGCATCGGCTGCGAC
AGATCCAGGAGAGACTCCACACAGCGGCGCTCCCAGCCTGAATCTGCCTGGATGGAAGTGA
GACCAATCATGCTGCAAGGAACACTTCCACGCCCCGTGAGGCCCTGTGCAGGGAGGAGCTG
CCTGTTCACTGGGATCAGCCAGGGCGCCGGGCCCCACTTCTGAGCACAGAGCAGAGACAGAC
GCAGGCGGGGACAAAGGCAGAGGATGTAGCCCCATTGGGGAGGGGTGGAGGAAGGACATGTA
CCCTTTCATGCCTACACACCCCTCATTAAAGCAGAGTCGTGGCATTTCAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

286/330

FIGURE 286

MPVPALCLLWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK
ARNSLGLYGRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK
VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ
IQERLHTAALPA

107153664

FIGURE 287

GGCAACATGGCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCAT
CCTGGTGATCACCTTACTCCTGGACCAGACCACCAGCCACACATCCAGATTA AAAAGCCAGGA
AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC
TGGACAGAAGTCAATGCCTTGAAGGAAATTCAAGCCCTGCAGACAGTCTGTCTCCGAGGCAC
TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCATGAGGCCAATG
AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC
CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA
CATGGTCACGGAAGGCAAGTTTGTGACGTCAACGGAATCGCTATCTCCTTCCTCAACTGGG
ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACGTGTGTCCTGTTCTCCCAATCAGCTCAG
GGCAAGTGGAGTGATGAGGCCTGTCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC
TAAATTAGGTCTTTCTCCAATGTGTCTCCTCCAAGCAAGATTCATCATAACTTATAGGTTTCATGA
TCTCTAAGATCAAGTAAAAATCATAATTTTTACTTATTAAAAAATTGCAACACAAGATCAAT
GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT
TCTTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT
AAACAGACTAAAATCTTTCTCTCTAGTCTTTCTCACTTGTACAAACCCAGTTTGTTTTCAA
AAATCACAGTAGCAATGCAACTCATCACTCTAGAAAAGCAAGCTTAGGCTACCTGAAAGATT
TTCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCCTACATCAGAGACTCTAGGT
GCTATATAATCCAAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC
TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT
GCCATATCAGAACACAAACCCCTGAAGAGGTTCTGATTTGATTTTTTTTTTTTCTTCATGCC
TACCCTTTTTTTTGAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT
CAATTTTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAACCCTAAGGCATAT
CAAAGAAGCAGATTGCATGATAAACGGAAATAGAAAAAAGAACCTACATTTATTTTGCTTT
AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCTTTTTTTTACATTT
TCGTATATTTATTTTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT
TGGAAGCTGAAAACGAATTTAAAGAATGCTATCTTGGAATTTGCATACGTCTGTGCAATT
TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT
TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTTAGAAGCAAACAATTTT
AAATATATTTTGTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT
CCCCTTTGCAAACTTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTTCATTGCTCA
ATAATAAAGCCTGAATTCTGATCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHKKRRVRDKDGLKTQIEKLWT
EVNALKEIQALQTVCLRGTKVHKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSQSAQ GK
WSDEACRSSKRYICEFTIPK

288/330

FIGURE 289

GCGAGGACCGGGTATAAGAAGCCTCGTGGCCTTGCCCGGGCAGCCGCAGGTTCCTCCGCGCGC
CCCGAGCCCCCGCGCC**ATGA**AAGCTCGCCGCCCTCCTGGGGCTCTGCGTGGCCCTGTCCTGCA
GCTCCGCTGCTGCTTTCTTAGTGGGCTCGGCCAAGCCTGTGGCCCAGCCTGTCGCTGCGCTG
GAGTCGGCGGCGGAGGCCGGGGCCGGGACCCTGGCCAACCCCCTCGGCACCCTCAACCCGCT
GAAGTCCTGCTGAGCAGCCTGGGCATCCCCGTGAACCACCTCATAGAGGGCTCCCAGAAGT
GTGTGGCTGAGCTGGGTCCCCAGGCCGTGGGGGCCGTGAAGGCCCTGAAGGCCCTGCTGGGG
GCCCTGACAGTGTTTGGCT**TGA**GCCGAGACTGGAGCATCTACACCTGAGGACAAGACGCTGCC
CACCCGCGAGGGCTGAAAACCCCGCCGCGGGGAGGACCGTCCATCCCCTTCCCCCGGCCCT
CTCAATAAACGTGGTTAAGAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAA

290/330

FIGURE 290

MKLAALLGLCVALSCSSAAFLVGS AKPVAQPVAALESAAEAGAGTLANPLGT LNPLKLLLS
SLGIPVNH LIEGSQKCV AELGPQAVGAVKALKALLGALT VFG

290/330

FIGURE 292

MKVVPSSL LLSVLLAQVWLV PGLAPSPQSPETPAPQNQTSRVVQAPREEEDEQEASEEKAGE
EEKAWLMASRQQLAKETS NFGFSLLRKISMRHDGNMVFSPFGMSLAMTGLMLGATGPTETQI
KRGLHLQALKPTKPGLLP SLFKGLRETLSRNLELGLSQGSFAFIHKDFDVKETFFNL SKRYF
DTECVPMNFRNASQAKRLMNHYINKETRGKIPKLFDEINPETKLILVDYILFKGKWLT PFDP
VFTEVDTFHLDKYKTIKVPMMYGAGKFASTFDKNFRCHVLKLPYQGNATMLVVLMEKMGDHL
ALEDYLT TDLVETWLRNMKTRNMEVFFPKFKLDQKYEMHELLRQMGIRRIFFSPFADLSELSA
TGRNLQVSRVLRRTVIEVDERGTEAVAGILSEITAYSMPPVIKVDRPFHFMIYEETSGMLLF
LGRVVNPTLL

292/330

FIGURE 293

CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCCAGAC**ATGAG**
GAGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTCCCAG
CACCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCAGAGAAG
GCCTGGGGCGCCCGTGTGGTGGAGCCTCCGGAGAAGGACGACCAGCTGGTGGTGCTGTTCCC
TGTCCAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTCAGGGCAGGGGGCCCCA
TCCTTCCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGCCGTGTCCTGAGTCCC
GAGCCCGACCATGACAGCCTGTACCACCCTCCGCCTGAGGAGGACCAGGGCGAGGAGAGGCC
CCGGTTGTGGGTGATGCCAAATCACCAGGTGCTCCTGGGACCGGAGGAAGACCAAGACCACA
TCTACCACCCCCAG**TAG**GGCTCCAGGGGCCATCACTGCCCCCGCCCTGTCCCAAGGCCCAGG
CTGTTGGGACTGGGACCCTCCCTACCCTGCCCCAGCTAGACAAATAAACCCCAGCAGGCAAA
AAAAAAAAAAAAAAAAA

294/330

FIGURE 294

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL
FPVQKPCLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE
RPRLWVMPNHQVLLGPEEDQDHIYHPQ

294/330

FIGURE 295

AGAAAAGCTGCACCTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGC
TACCCAAGGAAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGC
TGTTTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAA
TGGACCTGTTCTTCGTCTCCATCTCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC
TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTTATCTACCAGACCTTCT
GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATG
CGTGGGAAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCC
AGAGGGGGACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG
ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG
CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC
TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT
ATGGAGAAGGAAAGTGTGGACTGACAACGGCCCCGGTGATCCCTGTGGTCTATGATTTTGGC
GACGCCCAGAAAACAGCATCTTATTACTCACCTATGGCCAGCGGGAATTCAGTGCGGGATT
TGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGG
TCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGT
CCCCAGCAGTGTGGAGATTTTTCTGGTTTTGATTGGAGTGGATATGGAATCATGTTGGTTA
CAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGTTGAGAGTTTTTGTG
GGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCA
GTAGCTAGAATGTTAATGGCAGAAGAGAAAAACAATAAATCATATTGACTCAAGAAAAAAA

FIGURE 296

MNQLSFLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN
GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG
SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGL
YQKYPVKYGEKGCWTDNGPVI PVVYDFGDAQKTASYSPYGQREFTAGFVQFRVFNNERAAN
ALCAGMRVTGCNTEHHCIGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLL
FYR

FIGURE 297

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC
CACGAGGCTGCCGCATCCTGCCCTCGGAACA**ATG**GGACTCGGCGCGCGAGGTGCTTGGGCCG
CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCGCCCATGAAAGCGCAGCC
ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACCTCCAGTGCTAACTCAAC
AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAACTTCCAACAGTACTGTGAAAC
CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG
GCATCTAATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC
TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG
TAACCCACAATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACCTATGCAT
TCTGAAGCAAAGAAAGGATCAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC
GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC
GGTATCGAACCATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGA
ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAAACAATATTCT
CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA
AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT
TCATACAATGGTTTTAGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTGGCT
GGGGTGGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA
TGCCATCTGGGCATACAAATAAGAAGTTTGTACAGCACTCAGGATTTTGGGTATCTTTTGT
AGCTCACATAAAGAACTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA
CAGAAATTATACAATCAAACCTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG
TGCTTTAAACTGTAGTAGTTGGTCTAGAAACAAAATACTCC

FIGURE 298

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVP
SDH
TNETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSV
SQN
TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSI
LYIG
CKMYYSRRGIRYRTIDEHDAII

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FIGURE 299

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCTGAGCCCGAGCCGGGAGCCGG
TCGCGGGGGCTCCGGGCTGTGGGACCGCTGGGCCCCCAGCG**ATG**GCGACCCTGTGGGGAGGC
CTTCTTCGGCTTGGCTCCTTGCTCAGCCTGTCTGTCCTGGCGCTTTCCGTGCTGCTGCTGGC
GCAGCTGTCAGACGCCGCCAAGAATTTGAGGATGTCAGATGTAAATGTATCTGCCCTCCCT
ATAAAGAAAATTTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT
CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA
ATGCAAATATGAAGAAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA
TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCATACTGAAGAGG
CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT
TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG
AATATGCACAGCAGCGCTGGAAGCTTCAAGTCCAAGAGCAGCGAAAGTCTGTCTTTGACCGG
CATGTTGTCTCCTCAGC**TAA**ATTGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA
CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTTCATTTTAATACCTTGTTGATTTACCAACT
GTTGCTGGAAGATTCAAACTGGAAGCAAAACTTGCTTGATTTTTTTTTTCTTGTTAACGTA
ATAATAGAGACATTTTTTAAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCTATTTG
TGACTTTTACTAATAAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA
AGCACTCTCTTTTTTACCACATAGTTTTAACTTGACTTTCAAGATAATTTTCAGGGTTTTTG
TTGTTGTTGTTTTTTTGTGTTGTTTGGTGGGAGAGGGGAGGGATGCCTGGGAAGTGGTT
AACAACTTTTTTCAAGTCACTTTACTAAACAACTTTTGTAATAGACCTTACCTTCTATTT
TCGAGTTTTCATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG
ACTTTTGCCTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT
CTAAAATGCCTGGTGGCTTTTACAAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG
CAATGCATCCTAGAACAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG
GTGTGTGTGGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGGACACT
TGCAATAAAGAAATTTTATTTTAAACCCAAGCCTCCCTGGATTGATAATATATACATTTG
TCAGCATTTCCGGTCGTGGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT
AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT
TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAAATAATTTTTTGAAACATCAA

FIGURE 300

MATLWGGLRLGSLLSLSCLALSVLLLAQLSDAAKNFEDVRCKCICPPYKENS
QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIIIIYLSILG
VEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVILARSRSRANVLNKVEYA
QQRWKLQVQEQ
RKSVFDRHVLS

300/330

FIGURE 301

GCACCTGCGACCACCGTGAGCAGT**CATG**GCGTACTCCACAGTGCAGAGAGTCGCTCTGGCTT
CTGGGCTTGTCTGGCTCTGTCGCTGCTGCTGCCCAAGGCCTTCCTGTCCCGCGGGAAGCGG
CAGGAGCCGCGCCGACACCTGAAGGAAAATTGGGCCGATTTCCACCTATGATGCATCATCA
CCAGGCACCCTCAGATGGCCAGACTCCTGGGGCTCGTTTCCAGAGGTCTCACCTTGCCGAGG
CATTTGCAAAGGCCAAAGGATCAGGTGGAGGTGCTGGAGGAGGAGGTAGTGAAGAGGTCTG
ATGGGGCAGATTATTCCAATCTACGGTTTTGGGATTTTTTTATATATACTGTACATTCTATT
TAAGGTAAGTAGAATCATCCTAATCATATTACATCA**ATGA**AAATCTAATATGGCGATAAAAA
TCATTGTCTACATTAAACTTCTTATAGTTCATAAAATTATTTCAAATCCATCATCTCTTTA
AATCCTGCCTCCTCTTCATGAGGTACTTAGGATAGCCATTATTTTCAAGTTTCACATAAGAATG
TTTACTCAATGTTTAAGTGTTTTGCCCCAAAATTCACTAACAAGGCAGAACTAGGACTT
GAACATGGATCTTTTGGTTCTTAATCCAGTGAGTGATACAATTCAATGCACTCCCCTGCCA

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPPMMHHHQAPSDGQT
PGARFQRSHLAEAFKAKGSGGGAGGGGSGRGLMGQIIPIYGFGIFLYILYILFKVSRIILI
ILHQ

Downloaded from www.ashg.org

FIGURE 303

CGGCTCGAGTGCAGCTGTGGGGAGATTTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT
 GGATTTGAAAGTTGAGAGCAGC**ATG**TTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTAC
 TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG
 GGTGATTCACTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT
 AGACTGGACTCTGTCAACAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA
 ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC
 AATGATGGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA
 AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG
 AGGAGCCCCAAAGAGCTCATGGTCCATGTGGGTGGATTGATTGAGATGGGATGTGTTTTCCAG
 AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA
 GGAGATTGTATTTTCGTTACTACCACAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG
 GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATG
 CTTCAAGGAGTGAGGGAGTCAGATGGAGGAACTACACCTGCAGTATCCACCTAGGGAACCT
 GGTGTTCAAGAAAACCATTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCC
 CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC
 TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA
 GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG
 AAAAACCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCAATAATTGTA
 CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA
 CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACCTCACTTGAAAAAAGTCAGGTGGGG
 GAATGCCAAAAACACAGCAAGCCTTT**TGA**GAAGAATGGAGAGTCCCTTCATCTCAGCAGCGG
 TGGAGACTCTCTCCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTCAGACTCCCGCTCTC
 CCAGCTGTCCTCCTGTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG
 CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC
 CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCCGTT
 GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA
 AAAACCAACCCAAATCAA

FIGURE 304

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKLRMSVEYSQSWGHEFQNRVNLVGDI FRNDGSIMLQGVRES DGGNYTCSIHLGNLVFKKTIV
LHVSPEEPRTLVT PAALRPLVLGGNQLVIIVGIVCATILLLPVLILIVKKT CGNKSSVNSTV
LVKNTKKTNPEIKEKPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLR
SDRNNNSLEKKSGGGMPKTQQAF

304/330

FIGURE 305

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG
GTTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGG
AAACTACATTTTGC AAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATGCA**
GGATGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTG
GCCCTGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG
ATGGTTGTCTGGGCTGGTGGCTCTGGGGATTTGGTCTGTGTCATGCAGCGCAATTACCTACAAGA
TGAGAATGAAAATCGCACAGGAACCTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG
TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCCTGTGACACAAAC
TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACCTAACATGGGAAGAGAG
TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG
AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTCTGGATTATCTCGCCAGAAGTCGAAT
GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGAAGA
TGGAAAAGGAAATATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTG
AGAACAAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT
TAATGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT
ATGAATGCATCAGTAGCTGAAAAAAAAAAAAA

FIGURE 306

MQDEDGYITLNIKTRKPALVSVGPASSSWVRMALILLILCVGMVVGLVALGIWSVMQRNYL
QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSYGFRRHNLWE
ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEEFL
EDGKGNMNCAYFHNGKMHPTFCENKHVLMCERKAGMTKVDQLP

FIGURE 307

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG
CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGC
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG
CGGAGGAGAAGGAGGAGGAGGCGAACCAGAGAGGGGCGAGCAAAGAAGCGGTGGTGGTGGG
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTGAGCAGCCCCAGCAAAGGCAAGACCAGCTGC
GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT
CTGTTTAACTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACCTTTTCACACCTGAGTGCA
AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC
AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC
GTAGCCAGTGAGGGCAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGG
CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCKDNKLVFSRVKLFSGSKRRRRRP
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTDKEDSTYTLFNLI PVGLRVVAIQGVQTKLYLA
MNSEGILYLTSELTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMHNEST

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

FIGURE 309

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGG
 ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGG
 CAACCTGGATATTCTGAGACATATTTTGGGGGGATTTCAGTGAAAAAAGTGGGGGATCCCCCT
 CCATTTAGAGTGTAGCAAAGGAAAAAACACCAAGGTTGGGTTCCCTTCCTGACATTGGCAGTG
 CCCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCACACCCCTGTAGATTAC
 AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAAGGTGCCTGAAGATATTTAA
 ACCACGTCTTGGAATTTAGTGGGTCTTGGCTTTGGGATAGGTGAAGTGAGGACAGACACTG
 GAGAGGAGGGAAAGGGGACGTTTTCAATAGGAGGCAAACTCGAGGGTGGGATCCACTGAGG
 AGTACATAGGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCT
 GTGGAGGGGGGTACGTGAGGGGGGGGTCTGGGGCTTATCCTCAGGTCCTGTGGGTGGGGCAG
 CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG
 CGCGCTCCGGGCGCCTGCCGTTTGGGGGTGTCTCCTCCCGGGGCGCT**ATG**GCGGCGCTGGC
 CAGTAGCCTGATCCGGCAGAAGCGGGAGGTCCGCGAGCCCGGGGGCAGCCGGCCGGTGTGCG
 CGCAGCGGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG
 CTGTCCAAGGTGCGACTGTGCGGGGGGCGGCCCGCGCGGCCGGACCGCGGCCCGGAGCCTCA
 GCTCAAAGGCATCGTCACCAAAGTGTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCG
 ACGGAAGCATCCAGGGCACCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCT
 GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC
 TGAGGGACTGCTCTACAGTTCGCCGCATTTACAGCTGAGTGTGCTTTAAGGAGTGTGTCT
 TTGAGAATTACTACGTCCTGTACGCCTCTGCTCTCTACCGCCAGCGTCGTTCTGGCCGGGCC
 TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAA
 GGCAGCTGCCCACTTTCTGCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC
 ACAGTGTCCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCCC**TGA**AATGTAGTCCCTGGACTG
 GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACAACCTGT

310/330

FIGURE 310

MAALASSLIRQKREVPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR
GPEPQLKGIVTKLFCRQGFYLANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY
MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLEDKEGQVMKGNR
VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSPPAP

Tyrosine kinase phosphorylation site:

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

310/330

FIGURE 311

ATGGCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTG
GGACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGCTCTGCAACG
GCAACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGG
CGCCAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTT
GCAAATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTACACTCT
TCAACCTCATACCAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTAT
ATAGCCATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTACCCCTGAATGCAAGTT
TAAAGAATCTGTTTTTGAAAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAGG
AATCTGGTAGAGCCTGGTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGA
GTAAAGAAAACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCG
AGAACCATCTTTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAA
GCACAAGTGCGTCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACA**TAG**

FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWD RPSASRRRSSPSKNRGLCNGNLVDIFSKVRI FGLKKRRLR
RQDPQLKGIVTRLYCRQGYLQ MHPDGALDGTKDDSTNSTLFNLI PVGLRVVAIQGVKTGLY
IAMNGEGYLYPSELFTPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETV PKPGVTPSKSTSASAIMNGGKPVNKS KTT

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGI STNTRLLNLHENQIQI IKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYNLAMCNLREIPNLTPLIKLDLDELDSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLHNPWNCNCDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTNLNFTNVTVQDTGMYTCMVSNSVGN
TTASATLNVTAATTTTFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTKFTFTIPVTDINSIGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419,
434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
391-397, 422-428, 433-439, 531-537

FIGURE 315

GCGCCGGGAGCCCATCTGCCCCCAGGGGACAGGGGCGCGGGGCGGGCTCCCGCCCGGCACAT
 GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCCGTGCGGA
 GCGCCCGGGCGCCCGGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGGTCCGGGGATC
 GGG**ATG**TCCCTCCTCCTTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA
 CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACCTTTGCCCTGCCACCATCAACTGGGGC
 TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
 GTGGTGATCACTTACTCCAGTCGTATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
 AGTGGCCTTTGCTTCCAATTTCTGGCAGGAGATGCCTCCTTGACAGATTGAACCTCTGAAGC
 CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCAT
 GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
 AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTTGTGTATT
 ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
 GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA
 CCAGTGACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT
 ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
 ATTTTCCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
 GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
 CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT
 AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
 CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
 CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
 ACGGTCT**TGA**ATTACAATGGACTTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTC
 TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
 GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCCTTATACAATACCAAACAAGCAAA
 AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
 AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
 AGGTGAATATACCTAAAACCTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT
 TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACCTATTGGATT
 ATTAGTTATTTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
 TGAGCTAACCACTTCTAAGAAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
 TTCATTTGTGATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
 AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC
 TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT
 ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT
 CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
 TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
 CCCAACATACCATTTATAGTCTCTTCTTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
 GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
 TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
 ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

316/330

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDTLTQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

Downloaded from www.sciencedirect.com

FIGURE 317

CGCGAGGCGCGGGGAGCCTGGGACCAGGAGCGAGAGCCGCCTACCTGCAGCCGCCGCCACGGCACGGCAGCCA
 CC**ATG**GCGCTCCTGCTGTGCTTCGTGCTCCTGTGCGGAGTAGTGGAATTCGCCAGAAATTTGAGTATCACTACT
 CCTGAAGAGATGATTGAAAAAGCCAAAGGGGAACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGA
 CCAGGGACCGCTGGACATCGAGTGGCTGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTTAT
 ATCTGGAGACAAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAGTAATGATCTC
 AAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACGTGTCAGATATTGGCACATATCAGTGCAAAGTGAA
 AAAAGCTCCTGGTGTTGCAAATAAGAAGATTCATCTGGTAGTTCTTGTTAAGCCTTCAGGTGCGAGATGTTACG
 TTGATGGATCTGAAGAAATTGGAAGTGACTTTAAGATAAAATGTGAACCAAAGAAGGTTCACTTCCATTACAG
 TATGATGGCAAAAATTTGCTGACTCACAGAAAATGCCACTTCATGGTTAGCAGAAATGACTTCATCTGTTAT
 ATCTGTAAAAAATGCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAGAAAACAGAGTGGGCTCTGATC
 AGTGCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAAAGCTGGACTAATTGCAGGAGCCATTATAGGAACT
 TTGCTTGCTCTAGCGCTCATTGGTCTTATCATCTTTTGCTGTCGTAAGGAGCGCAGAGAAGAAAAATATGAAAA
 GGAAGTTCATCACGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACATCG
 GCAGTAATCATTCATCCCTGGGGTCCATGTCTCCTTCCAACATGGAAGGATATTCCAAGACTCAGTATAACCAA
 GTACCAAGTGAAGACTTTGAACGCACCTCTCAGAGTCCGACTCTCCACCTGCTAAGTTCAAGTACCTTACAA
 GACTGATGGAATTACAGTTGTAT**TAA**ATATGGACTACTGAAGAATCTGAAGTATTGTATTATTTGACTTTATTTT
 AGGCCTCTAGTAAAGACTTAAATGTTTTTTAAAAAAGCACAGGCACAGAGATTAGACGAGCTGTAAAGAACAC
 ATCTACTTTTATGCAATGGCATTAGACATGTAAGTCAGATGTGATGTCAAATTAGTACGAGCCAAATTCCTTGT
 TAAAAACCCCTATGTATAGTGACACTGATAGTTAAAGATGTTTTATTATATTTTCAATAACTACCACTAACAA
 ATTTTTAACTTTTCATATGCATATTCTGATATGTGGTCTTTTAGGAAAAAGTATGGTTAATAGTTGATTTTTCAA
 AGGAAATTTTAAATTCCTTACGTTCTGTTTAAATGTTTTGCTATTTAGTTAAATACATTGAAGGGAAATACCCG
 TTCTTTTCCCCTTTTATGCACACAACAGAAACACGCGTTGTCATGCCTCAAACCTATTTTATTTGCAACTACA
 TGATTTTACACAATTCCTTAAACAACGACATAAAATAGATTTCTTGTATATAAATAACTTACATACGCTCCA
 TAAAGTAAATTCCTCAAAGGTGCTAGAACAAATCGTCCACTTCTACAGTGTCTCGTATCCAACAGAGTTGATGC
 ACAATATATAAATACTCAAGTCCAATATTAAGAAAGATTATTTCTCAATAACTTCTATAAAAAATAAGTTTGATGG
 TTTGGCCCATCTAACTTCACTACTATTAGTAAGAACTTTTAACTTTTAAATGTGTAGTAAGGTTTATTCTACCTT
 TTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTGAGGTGCTAACATGTGAGGATTAATCCAGTGAT
 TCCGGTCACAATGCATTCAGGAGGAGGTACCCATGTCACTGGAATTGGGCGATATGGTTTATTTTCTTCCC
 TGATTTGGATAACCAAATGGAACAGGAGGAGGATAGTGATTCTGATGGCCATTCCCTCGATACATTCTGGCTT
 TTTTCTGGGCAAAGGGTGCCACATTGGAAGAGGTGGAATATAAGTTCTGAAATCTGTAGGGAAGAGAACACAT
 TAAGTTAATTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCTCATTAAAGACAAAGTTACCCACAACACT
 GAGATCACATCTAAGTGACACTCCTATTGTGAGGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAA
 TGTATAACAGGTGACCAATGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCTTAAACAA
 CTTCAACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAGGACATGCTTGTTTTAGTCCAGTGGTTT
 CCACAGCTGGCTAAGCCAGGAGTCACCTGGAGGCTTTTAAATACAAAACATTGGAGCTGGAGGCCATTATCCTT
 AGCAAACCTAATGCAGAAACAGAAAATCACTACCGCATGTTCTCATTATAAGTGGGAGGTAATGATAAGAACT
 TATGAACACAAAGAAGGAAACAATAGACATTTGGAGTCTATTTGAGAGGGGAGGGTGGGAGAAGGAAAAAGGAGCA
 GAAAAGATAACTATTGAGTACTGCCTTCACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGTGACACA
 TGTTTACCTATGGAACAAACCTTCATGTGTATCCCTAAACCTAAAAATAAAAGTTAAAAAARAAAAA
 AA
 AAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 318

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFARSL SITTP EEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK
APGVANKKIHVLVVLVKPSGARCIVDGS E EIGSDFKIKCEPKESLPLQYEWQKLSDSQKMPT
SWLAEMTSSVISVKNASSEYSGTY SCTVRNRVGS DQCLLR LNVVPPSNKAGLIAGAIIGTLL
ALALIGLIIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM
EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

Signal sequence.

amino acids 1-19

Transmembrane domain:

amino acids 236-257

N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

N-myristoylation sites.

amino acids 116-122, 208-214, 219-225, 237-243, 241-247,
245-251, 296-302

Myelin P0 protein.

amino acids 96-125

318/330

FIGURE 319

TGAAATGACTTCCACCGGCTGGGACGGGAACCTTCCACCCACAGCTATGCCTCTGATTGGTGTA
ATGGGTGAAGGTGCCTGTCTAACTTTTTCTGTAAAAAGAACCAGCTGCCTCCAGGCAGCCAGCC
CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTTCG
CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG
AGAGATGAATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC
CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG
CTTCTCTGGAGCCAGGTATCAGGGGGCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGT
GAAGGGGGTTGTTCCCCAGAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG
CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGAT
GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACGTGTTTTCAAAAA
CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAAC
TTGTTCTCATCGTGTCAAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC
AGTGCACACAGGCGGTTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC
TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC
TCTTGAATGTCTAGACCAGGACCTCCCTCCCCCTGGCACTGGTTTGTTCCTGTGTCATTTCA
AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATT
TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG
GTGCCTCTGGATGCTGTGAAGAGTCTACAGAGAAGATTCTTGTATTTATTACAACTCTATTT
AATTAATGTGAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG
CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG
TGGATGGGTGCTTAGTAAGTACTTAATAAACTGTGGTGCTTTTTTTGGCCTGTCTTTGGATT
GTTAAAAAACAGAGAGGGATGCTTGGATGTAAACTGAACTTCAGAGCATGAAAATCACACT
GTCTTCTGATATCTGCAGGGACAGAGCATTGGGGTGGGGTAAGGTGCATCTGTTTGAAAAG
TAAACGATAAAATGTGGATTAAAGTGCCAGCACAAAGCAGATCCTCAATAAACATTTTATT
TCCCACCCACACTCGCCAGCTCACCCCATCATCCCTTTCCCTTGGTGCCCTCCTTTTTTTTT
TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC
ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCCTGCTAATAAAAGACAACATAA
CTCCAAAAA

320/330

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFPCPPLLATASQMOMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK
GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLEFYLKTVFKNH
HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL
TKALGEVDILLTWMQKFYKL

Signal sequence:

amino acids 1-42

N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

320/330

FIGURE 321

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGCATGAAGTTACAGTGTGTTTCCCTTTGGCTC
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTC
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAAGAGCCATCCAAGCTAAGG
ACACCTTCCCAAATGTCACATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAAGGATCA
TCAGGAGCCAAACCCCAAAATCCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA
TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC
AATGCCACCAGAGTCATCCATGACAACTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
ATCCCTGGGAGAGCTCGACGTCTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
CAGCTTGATGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT
GTGGGAGACAGCCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGCAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGCAAAAAGTCTACTGTGGTAT
TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCCATATTTTACCTATGA

FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282

><subunit 1 of 1, 177 aa, 1 stop

><MW: 20452, pI: 8.00, NX(S/T): 2

MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETIQI IKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ
RQCHCRQEATNATRVIHNDYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

FIGURE 323

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG
AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT
AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACCTGCACCTC
GGTTCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACG**ATG**TTGGGGGCCCCGCCT
CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA
ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC
AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC
CATCTACAGTGGCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA
TGAGCAGAAGATACCTCTGCATGGATTTTTCAGAGGCAACATTTTTGGATCACACTATTTTCGAC
CCGGAGAAGTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC
TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC
CACCCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACCTCAACACC
CCCATACCACGGCGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGT
GCTGAAGCCCCGGGCCCCGGATGACCCCGGCCCCGGCCTCCTGTTTACAGGAGCTCCCGAGCG
CCGAGGACAACAGCCCAGTGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTCGAGTGAAC
ACGCACGCTGGGGGAACGGGCCCCGAAGGCTGCCGCCCTTCGCCAAGTTCATC**TAG**GGTTCG
CTGG

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSWSGGLIHLYTATARN SYHLQIHKN GHVD
GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGY
DVIHSPQYHFLVSLGRAKRAFLPGMNPPPYSQLSRRNEIPLIHFNTPIPRRHTRSAEDDSE
RDPLNVLKPRARMT PAPASCSQELPSAEDNSPMASDPLGVVRGGRVNT HAGGTGPEGCRPFA
KFI

Important features of the protein:**Signal peptide:**

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

FIGURE 325

GGAAAAAGGTACCCGCGAGAGACAGCCAGCAGGACTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG**
GGCTGTCTCTGGGGTCTGGGCTCTGCCCCCTTTTCTTCTTCTGCTGGGAGGTTGGGGTCTCTGG
GAGCTCTGCAGGCCCCAGCACCCGCGAGAGCAGACACTGCGATGACAACGGACGACACAGAAG
TGCCCGCTATGACTCTAGCACCCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG
ACCTCTTCTAGGGCCTCAACCCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA
GAGAATTTCCCTGCAAGAGAGACCAGGAGTTTCAAAAAACATCTCCCAACTTCATGGTGC
TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC
ACAGTTCAGACCATCACAGGCAGTGATCCCAGGAAGCCATCTTTGACACCCTTTGCACCGA
TGACAGCTCTGAAGAGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA
CAGAAGCTAAGGGCCTGTCTCAGAGAGCAGTGCCTCTTCCGACGGCCCCCATCCAGTCATC
ACCCCGTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC
GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCCTCAT
GGTCCCCGGGATCTGATGTCACTCTCCTCGCTGAAGCCCTGGTGACTGTCACAAACATCGAG
GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA
CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC
CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC
ACAGCCGGCACCCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA
CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA
CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC
AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTGAGCAGTGGGCAAAACAACTTC
CTTTGCTGGGAGCTCTGCTTCTCCTACAGCCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC
CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCCT
CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA
GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCCACGACTGCCCGGAC
GAGGCCGACCACAGACG**TGA**GTGCAGGTGAAAATGGAGGTTTCTCTCTCTGCGGCTGAGTG
TGGCTTCCCCGGAAGACCTCACTGACCCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC
CGGGAACCTCCACGCCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA
ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCTTAGCCTG
GGCCCCCACCAGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG
AAGGGCAGCATGTCCAAGCCCCTAACCCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC
CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCCAGAGGTGTCTTGGA

FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTL
SAETSSRASTPAGPIPEAETRGAKRISPARETRSF TKTSPNFMVLIATSVETSAASGSPEGAGM
TTVQTITGSDPEEAI FDTLCTDDSS EEA KT LTMDILT LAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLAEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTP LPTNSATEREVTAPGATTLSGALVTVSRNPLEETSALSVETPSY
VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNF TPSETPTMDIATKGPFPTSRD
PLPSVPPTTTNSSRGTNSTLAKITTS AKTTMKPQQPRPRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-90

Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169,
246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339,
404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-
320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

CGCGAGCATCCGCTGCGGTCCTCGCCGAGACCCCGCGGATTCGCGGTCCTTCCCGCGG
GCGCGACAGAGCTGTCTCGCACCTGGATGGCAGCAGGGGCGCCGGGGTCTCTCGACGCCA
GAGAGAAATCTCATCATCTGTGCAGCCTTCTTAAAGCAAACCTAAGACCAGAGGGAGGATTAT
CCTTGACCTTTGAAGACCAAACTAACTGAAATTTAAAATGTTCTTCGGGGGAGAAGGGAG
CTTGACTTACACTTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT
GCCTCAAAAAGAGTCTAGAAGATGTTGTCAATTGACATCCAGTCATCTCTTCTAAGGGGAATC
AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAAC
AAAAAACATATCAGGGGACAAAGCATGTAACCTGATGATCTTCGACACTCGAAAAACAGCTA
GACAACCCAACTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA
AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG
CCAAGAGTTACCCAGGAAGATTCTCTTACATGGCCAATTTTCACAAGCAGTCACTCCCC
TAGCCCATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTTCT
CAGAAGTTTGGATCCTCAGATCACCTGGAGAACTATTTAAGATGGATGAAGCAAGTGCCCA
GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTCACAATTTTCTCTGATCAAGAAA
TAGCTCATCTGCTGCCTGAAAATGTGAGTGCCTCCAGCTACGGTGGCAGTTGCTTCTCCA
CATACCACCTCGGCTACTCCAAAGGCCACCCTTCTACCACCAATGCTTCAGTGACACC
TTCTGGGACTTCCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACTTCTC
AGCCTCCCACGACCCTCATTTCTACAGTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG
GCTACAACAGCAGTTCTGACTACCACCTTTCAGGCACCTACGGACTCGAAAGGCAGCTTAGA
AACCATAACGTTTACAGAAATCTCCAACCTTAACCTTTGAACACAGGGAATGTGTATAACCCTA
CTGCACTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAAACTGCTTCCTGGGAAGGT
AGGGAGGCCAGTCCAGGCAGTTCTCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC
ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCTGTTCTTGGTGATAGGCC
TCGTCTCTGCGGTAGAAATCCTTTCGGAATCACTCCGAGGAAACGTTACTCAAGACTGGAT
TATTTGATCAATGGGATCTATGTGGACATCTAAAGGATGGAACCTCGGTGTCTCTTAATTCATT
TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG
TATTTTGAAGACAGGAAAATGCCCCCTTCTGCTTTCTTTTCTTTTGGAGACAGAGTCTT
GCTCTGTTGCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCAGCAACCTCCGTCTC
CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA
CCACACCTGGGTGATTTTTGTATTTTTAGTAGAGACGGGGTTTACCATGTTGGTCAGGCTG
GTCTCAAACCTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG
CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTGGTTTTTGGAGAAGGAATGAAGTG
GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAAATATTAGCTAAAAACAAAGCTCT
ATGTAAAGTAATAAAGTATAATTGCCATATAAATTTCAAATTCAACTGGCTTTTATGCAAA
GAAACAGGTTAGGACATCTAGGTTCCAATTCATTACATTCTTGGTTCCAGATAAAATCAAC
TGTTTTATATCAATTTCTAATGGATTTGCTTTTCTTTTTATATGGATTCTTTTAAACTTAT
CCAGATGTAGTTCTTCCAATTAAATATTTGAATAAATCTTTTGTTACTCAA

FIGURE 328

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLKGIKRGNEPVYTSTQED
CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF
KMDEASAQLLAYKEKGHSQSSQFSSDQEIHAHLLPENVSALPATVAVASPHHTSATPKPATLL
PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAATLQAMATTAVLTTTFQAP
TDSKGSLETIPFTEISNLTNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV
PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSLDYLINGIYVDI

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

FIGURE 329

CTCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
 CCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
 CCCTGCAGTGACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
 GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
 GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA
 ACCTCACCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
 TCTTTACTGATCTCTCTGTTCTTTCCAGGACCCTGCTGTCCTCCCTCCCCTTCTCCCAC
 CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC
 CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
 GCTGAGGGCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
 AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCCATGCAGC
 TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
 GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGGTGCTGCTGAGCCTTCTGTGAGC
 CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
 CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
 CCTTCCCAGGGCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA
 GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGGCCAGCAGT
 GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCACCTCAGCCTCAGAG
 TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGATGTTCCA
 GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
 GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA
 GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCTGGGC
 CTCATGCCCAGTGTGCGACCCTGCCTTCCCTCCCACTCCAGACCCACCTTGTCTTCCCTCCC
 TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT
 GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
 GTGAAAAACGTGATTCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG
 GACTCTGAATTCTAACAATGCCCAGTGACTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTG
 ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
 CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG
 TCCAGGCCTTGGTCAGGTGAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG
 TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCTTTTGAAAAAATGATGAAGA
 AAACCTTGGCTCCTTCCTTGTCTGGAAAGGGTACTTGCCTATGGGTCTGGTGGCTAGAGA
 GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
 ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA
 GCACAACCTACTATTTTTTTTTCTTTTTCCATTATTATTGTTTTTAAGACAGAATCTCGTGCT
 GCTGCCCAGGCTGGAGTGCAGTGGCAGCATCTGCAAACCTCCGCCTCCTGGGTCAAGTGATT
 CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
 TTTGTACTTTTAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGAC
 CTCAAATGAGCCTCCTGCTTCACTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
 TCTGGCCCTATTTCTTTTAAAAAGTGAAATTAAGAGTTGTTGAGTATGCAAACTTGGAAG
 ATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT
 TATTTGTTTTGTTGTAATTTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTCTT
 TTTACAGAGCAATTATCTTGATATACAACCTTGTATCCTGCCTTTTCCACCTTATCGTTCC
 ATCACTTTATTCAGCACTTCTCTGTGTTTTACAGACCTTTTTATAAATAAAATGTTTCATCA
 GCTGCATAAAAAAAAAAAAAA

FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDSELLISLFV
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSI PMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEKEAPSQAPEGD
VISMPPLHTSEEEELGFSKFVSA

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128